METHODS AND COMPOSITIONS FOR MODULATING FERTILITY

This invention was made in part with government support under Grant numbers HD U54 29099 and P30 28934 awarded by the National Institutes of Health. The United States government has certain rights in the invention.

1. INTRODUCTION

The present invention relates to vertebrate FSP95 genes and their encoded protein products, as well as derivatives and analogs thereof. Production of vertebrate FSP95 proteins, derivatives and antibodies is also provided. The invention further relates to therapeutic compositions and methods of diagnosis and therapy, including compositions and methods for modulating fertility (*e.g.*, inhibiting or promoting), such as methods for contraception, and methods for promoting fertility.

2. BACKGROUND OF THE INVENTION

Once deposited in the female reproductive tract, ejaculated mammalian spermatozoa undergo an intricate series of membrane and metabolic changes, collectively termed capacitation, which prepare the sperm to undergo the acrosome reaction and to bind and penetrate the zona pellucida (Yanagimachi, 1994, in: *The Physiology of Reproduction* (Knobil, E., and Neill, J. D., eds) pp. 189-3 17, Raven Press, New York, NY). During capacitation of bull and human spermatozoa, an increase in intracellular Ca^{++} (Handrow et al., 1989, J. Exp. Zool. 252:174-182, Baldi et al., 1991, J. Androl. 12:323-330) and cyclic 3',5'-adenosine monophosphate (cAMP) has been detected (Parmnaud and Milhet, 1996, J. Clin. Endrocrinol. Metab. 81:1357-1360, Parrish et al., 1994, Biol. Reprod. 51:1099-1108). Capacitated spermatozoa are also characterized by a rise in intracellular pH (Vredenburgh-Wilberg and Parrish, 1995, Mol. Reprod. Dev. 40:490-502), and loss of cholesterol from the sperm plasma membrane (Visconti and Kopf, 1998, Biol. Reprod. 59:1-6).

Protein tyrosine phosphorylation has also been associated with capacitation, as well as motility of mammalian sperm. During capacitation, human spermatozoa undergo tyrosine phosphorylation of a characteristic set of proteins; however, thus far only one such protein has been cloned and localized (Turner et al., 1998, J. Biol. Chem. 273:32135-32141). Phosphorylation of tyrosine residues on a cohort of specific proteins in mouse (Visconti et al., 1995, Development 121:1129-1137), bovine (Galantino-Homer, et al., 1997, Biol. Reprod. 56:707-7 19) and human sperm (Carrera et al., 1996, Dev. Biol. 180:284-296, Luconi et al., 1996, Biol. Reprod. 55:207-216, Leclerc et al., 1998, Androl. 19:434-443) indicates that activation of sperm tyrosine kinases occurs during the

capacitation process. Mouse caput epididymal sperm which lack the ability to undergo capacitation do not display tyrosine phosphorylation of this set of proteins (Visconti et al., 1995, Development 121:1129-1137).

Sperm acquire the capacity for motility during epididymal maturation, become motile in the ejaculate, and express a distinct motility pattern called hyperactivation during capacitation (Yanagimachi, 1994, in: *The Physiology of Reproduction* (Knobil, E., and Neill, J. D., eds) pp. 189-3 17, Raven Press, New York, NY). The initiation and maintenance of motility of spermatozoa involves phosphorylation-dephosphorylation events affecting the activities of protein kinase substrates (Tash, 1989, Cell Motil. Cytoskeleton 14:332-339, Tash and Bracho, 1994, J. Androl. 15:505-509).

Many sperm proteins are potent auto- and iso-antigens which evoke immune responses in both males and females, and antisperm antibodies ("ASA") are capable of causing infertility in both human and animal models (Primakoff et al., 1988, Nature 335:543-546, Kutteh et al., 1996, Am. J. Reprod. Immunol. 35:429-433, Ohl and Naz, 1995, Urology 46:591-602, Nip et al., 1995, Hum. Reprod 10:564-2469). For example, ASA can be measured in serum of up to 70 percent of men after vasectomy compared with 2 to 8 percent of men without vasectomy (Raspa, 1993, Am. Fain. Physician 48:1264-1268). The incidence of ASA in infertile couples varies from 9 to 36% and in infertile men is from 8 to 21% (Gubin et al., 1998, Am. J. Reprod. Immunol. 39:157-160), while the prevalence of ASA in the general population ranges from 0 to 2% (Jarow and Sanzone, 1992, J. Urol. 148:1805-1807). The presence of ASA in the female genital tract can affect sperm-egg interaction by several mechanisms (D'Cruz et al., 1991, J. Immunol. 146:611-620, Peters and Coulam, 1992, Am. J. Reprod. Immunol. 27:156-162, Menge, 1971, Proc. J. Exp. Biol. Med. 138:98-102, Clarke et al., 1986, Fertil. Steril. 46:435-441), and consequently can result in reduced fertility (Ayvaliotis et al., 1985, Fertil. Steril. 43:739-742, Mandelbaum et al., 1987, Fertil. Steril. 47:644-651, Kobayashi et al., 1990, Fertil. Steril. 54:1107-1113). Accordingly, there is a need in the art for novel methods and compositions for modulating fertility.

Citation of references in this section or in any other section shall not be construed as an admission that such references are prior art to the present invention.

3. SUMMARY OF THE INVENTION

The present invention is based, in part, on the discovery of the nucleotide and amino acid sequence of a novel fibrous sheath protein of 95 kDa (FSP95), which is localized to the ribs of the fibrous sheath in the principal piece of sperm tail, undergoes tyrosine phosphorylation during capacitation of human spermatozoa and plays a role in

sperm motility. The 853 amino acid residue protein has a calculated molecular weight of 94.6 kDa and a pI of 6.0 and contains multiple potential phosphorylation sites for protein kinase C and casein kinase II as well as one tyrosine kinase phosphorylation site at amino acid position 435.

The present invention is directed to nucleotide sequences of vertebrate fibrous sheath protein of 95 kDa (FSP95), and amino acid sequences of their encoded proteins, as well as derivatives (*e.g.*, fragments) and analogs thereof. Nucleic acids hybridizable to or complementary to the FSP95 nucleotide sequence are also provided. In a specific embodiment, the FSP95 protein is a mammalian protein, preferably a human protein.

The invention is directed to vertebrate FSP95 derivatives and analogs which are functionally active, *i.e.*, capable of displaying one or more known functional activities associated with a full-length (wild-type) FSP95 protein. Such functional activities include but are not limited to antigenicity [ability to bind (or compete with FSP95 for binding) to an anti-FSP95 antibody], immunogenicity (ability to generate antibody which binds to FSP95).

The invention is further directed to fragments (and derivatives and analogs thereof) of a vertebrate FSP95 protein that comprise one or more domains of the FSP95 protein, including but not limited to the extracellular domain, transmembrane domain, intracellular domain, intracellular targeting domain, RII binding site, or any combination of the foregoing. In specific embodiments, fragments of FSP95 comprise amino acids 318-335, amino acids 368-385, and/or amino acids 671-688.

Antibodies to a vertebrate FSP95 protein, its derivatives and analogs are additionally provided by the present invention.

Methods for production of vertebrate FSP95 proteins, derivatives and analogs, *e.g.*, by recombinant means, are also provided.

The present invention is also directed to therapeutic and diagnostic methods and compositions based on vertebrate FSP95 proteins and nucleic acids. The invention provides for treatment of sperm motility disorders by administration of a therapeutic compound of the invention. Such therapeutic compounds include, but are not limited to, FSP95 proteins and analogs and derivatives (including fragments) thereof; antibodies thereto; nucleic acids encoding the FSP95 proteins, analogs or derivatives; and FSP95 antisense nucleic acids. In a preferred embodiment, a therapeutic of the invention is administered to treat infertility. In other specific embodiments, a therapeutic of the invention is administered to decrease sperm motility. In yet other specific embodiments, a therapeutic of the invention is administered to increase sperm motility.

In one embodiment, therapeutics which antagonize, or inhibit, FSP95 function (referred to as "Antagonist Therapeutics") are administered for therapeutic effect. In another embodiment, therapeutics which promote FSP95 function (referred to as "Agonist Therapeutics") are administered for therapeutic effect.

Disorders of sperm motility, involving aberrant or undesirable levels of expression or activity or localization of FSP95 protein can be diagnosed by detecting such levels, as described more fully *infra*.

In a preferred aspect, a therapeutic of the invention is a fragment of an antibody to FSP95 consisting of at least the binding domain of the antibody. In yet another preferred aspect, a therapeutic of the invention is a fragment of FSP95 comprising amino acids 318-335, 368-385 and/or 671-688.

3.1 DEFINITIONS

Abbreviations used herein are: cAMP, cyclic 3',5'-adenosine monophosphate; AKAP, A-kinase anchor protein; PK-A, protein kinase A; ASA, antisperm antibodies; FSP95, fibrous sheath protein 95 kDa; OBG, octyl-3-D-glucopyranoside; BSA, bovine serum albumin; PBS, phosphate buffer saline; RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction; IPTG, isopropyl-1-thio-β-D-galactopyranoside; rFSP95, recombinant FSP95; HSA, human serum albumin; SIY, sperm immobilization value; pro-mAKAP82, precursor of mouse AKAP82; pro-hAKAP82, precursor of human AKAP82.

4. DESCRIPTION OF THE FIGURES

FIG.1. Identification of FSP95 as a high molecular weight major acidic tyrosine phosphorylated protein in capacitated human spermatozoa by 2-D immunoblot. Fresh swim up sperm were capacitated in absence and presence of genistein, a protein tyrosine kinase inhibitor. The location of FSP95 spot in the 2-D sperm proteome is indicated with a white half rectangular box (A) in a silver stained gel. The high molecular weight acidic proteins (box area of A) which showed phosphorylation of tyrosine residues are shown in blots probed with antiphosphotyrosine antibody before (B) and after capacitation (C). Prominent high molecular weight acidic tyrosine phosphorylated proteins (~95 kDa, pI 5.1-5.5) show increased phosphorylation with capacitation (C). A lack of phosphorylation of the 95 kDa proteins was observed in capacitated cells in presence of genistein (D). The spot which was cored for microsequencing FSP95 is indicated with a white circle along with an arrow(A).

FIG. 2. Antigenicity of FSP95 in men and women with antisperm antibodies. Human sperm proteins were separated by two dimensional SDS-PAGE, transferred to nitrocellulose membranes and probed with sera from an infertile male (A) and from an infertile female (B) previously screened for antisperm antibodies by the immunobead binding test. The immunoreactive proteins were compared to Western blots probed with sera from clinically fertile male (C) and female (D) subjects. Note the strong immunoreactivity exhibited by proteins at 95 kDa including FSP95 (arrow, the protein spot selected for microsequencing) using sera from infertile subjects of both sexes compared to only faint immunoreactivity observed in the 95 kDa group with fertile subjects sera.

FIG. 3. Nucleotide and deduced amino acid sequences of the human sperm protein FSP95. The deduced amino acid sequence of human sperm FSP95 is shown below the cDNA sequence. The numbers on the left refer to the nucleotide sequence; numbers on the right refer to the amino acid sequence. The consensus ATG of the open reading frame and the polyadenylation signal (ATFAAA) are indicated in bold letters. The termination codon (TAA) is marked with an asterisk. The 5-prime and 3-prime untranslated regions are 162 bp and 218 bp respectively and are shown in italics. The calculated molecular weight and p1 of the predicted protein were 94.6 kDa and 6.0 respectively. The 18 underlined sequences indicate the tryptic peptides obtained by microsequencing. The putative tyrosine kinase phosphorylation site is indicated in bold within a box (residue number 435). The nucleic acid sequence was submitted to the GenBank (accession number AF087003).

FIG. 4. Homology comparison of the deduced amino acid sequences of human sperm FSP95 with those of mouse and human sperm fibrous sheath AKAPs (mouse: pro-mAKAP82, accession # 148968; human: pro-hAKAP82, accession # AF072756). The sequences were listed in descending order of homology from the FSP95. The alignment was constructed by use of the GCG-PILEUP program and formatted with ALSCRIPT version 2.0. The shaded areas indicate the amino acid identities and similarities among the molecules (cut off 8 in ALSCRIPT). The conserved AKAP-like intracellular targeting domains are shown in boxes. The N-terminal RII-binding domain of the mouse and human AKAP82 is highlighted with an underline.

FIG. 5. Analysis of human sperm FSP95 expression. A: Northern blot (Clontech) containing 2 μg of poly(A)+ mRNA from human tissues per lane was hybridized with radiolabelled FSP95 cDNA and exposed for 60h. The migration of RNA markers (kb) are indicated on the left. A single transcript of - 3.0 kb was apparent only in testis. The blot was subsequently stripped and rehybridized with actin to assess the levels of RNA in each lane (data not shown). B: Dot blot containing poly(A)+ RNA from 50 human tissues (obtained from Clontech) was hybridized with a radiolabelled FSP95 cDNA and signals

were visualized by autoradiography. A hybridization was found only in testis following 18 h of exposure. The dot blot contained normalized amounts (89-514 ng) of Poly(A)+ RNA from the following 50 tissues: Al, whole brain; A2, amygdala; A3, caudate nucleus; A4, cerebellum; AS, cerebral cortex; A6, frontal lobe; A7, hippocampus; A8, medulla oblongata; B 1, occipital lobe; B2, putamen; B3, substantia nigra; B4, temporal lobe; BS, thalamus; B6, subthalamic nucleus; B7, spinal cord; Cl, heart; C2, aorta; C3, skeletal muscle; C4, colon; CS, bladder; C6, uterus; C7, prostate; C8, stomach; Dl, testis; D2, ovary; D3, pancreas; D4, pituitary gland; DS, adrenal gland; D6, thyroid gland; D7, salivary gland; D8, mammary gland; El, kidney; E2, liver; E3, small intestine; E4, spleen; ES, thymus; E6, peripheral leukocyte; E7, lymph node; E8, bone marrow; Fl, appendix; F2, lung; F3, trachea; F4, placenta; Gi, fetal brain; G2, fetal heart; G3, fetal kidney; G4, fetal liver; GS, fetal spleen; G6, fetal thymus; G7, fetal lung.

FIG. 6. Isolation of recombinant FSP95 (rFSP95) from E. coli and immunoblotting of recombinant and sperm FSP95. A portion of the human sperm FSP95 cDNA (encoding residues 1-779 of the protein) was expressed in E. coli using the pET2Sb plasmid, induced by addition of 1.0 mM IPTG and purified by nickel ion affinity column chromatography followed by preparative polyacrylamide gel electrophoresis. A: Bacterial extracts stained with Coomassie; lane 1, uninduced; lanes 2 and 3, 1.5 h and 3.0 h after induction; expressed FSP95 is indicated by the arrow. B: Coomassie stained purified rFSP95 (2.3 μg), used for immunization; left, molecular weight markers. C: Immunoblot of gel purified rFSP95 (3.5 μg) probed with rat antisera against rFSP95 at 1:SOOO dilution; lane 1, preimmune serum; lane 2, immune serum. D: 2-D Blot of human sperm proteins probed with rat serum against gel purified rFSP95 . The 95 kDa region at pI ~ 5.3 originally microsequenced immunoreacts with the antibody (arrow). The pH gradient is indicated at the top. Molecular weight markers are shown to the left of all the panels.

FIG. 7. Phase contrast (A, C) and indirect immunofluorescence staining (B, D) of FSP95 in capacitated permeabilized human spermatozoa. Immunofluoresence was noted throughout the principal piece (PP) of the flagellum (bar) with immune sera (B) while no fluorescence was observed in the head, mid piece (MP) or the end piece (EP). No immunostaining was observed with preimmune sera (D), or by using immune sera on live capacitated unpermeabilized sperm (data not shown).

FIG. 8. Electron microscopic immunogold localization of FSP95 in a longitudinal (A) and a cross section (B) of the principal piece of ejaculated human spermatozoa. Gold particles (arrows) were detected in the ribs of fibrous sheath (FS) but not within the central portion of the longitudinal columns (LC) . No immunoreactivity was

detected on the outer dense fibers (ODF) or on the axoneme (AXO). Only a rare gold particle was detected on control sections (C&D) treated with preimmune rat serum .

FIG. 9. Tyrosine phosphorylation of FSP95 during in vitro capacitation of human spermatozoa. Proteins of uncapacitated sperm (A, C) were compared to 6 h capacitated sperm (B, D) after 2D SDS -PAGE, electroblotting and probing with rat antisera against rFSP95 (A, B) and antiphosphotyrosine monoclonal antibody (C, D). Immunoreactive forms of FSP95 with p1 of - 5.3 (A) are less abundant after capacitation (B), while more acidic forms of FSP95 are immunoreactive after capacitation (B, arrows). These acidic charge shifts in immunoreactive forms of FSP95 after capacitation are accompanied by increased tyrosine phosphorylation and the appearance of more acidic phosphotyrosine containing isoforms (C, D, arrows).

FIG. 10. Helical wheel representation of FSP95 RII subunit binding domain alpha helix (amino acid residues 671-688), indicating surface of hydrophobic residues, shown in boxes (□).

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention provides "*FSP95*"(fibrous sheath protein of 95KDa) polynucleotides, polypeptides, and derivatives and analogs thereof. The invention further provides methods for the use of such sequences in contraception and modulation of fertility.

Production of recombinant FSP95 proteins, and peptide fragments comprising functional domains, derivatives, antibodies, and ligands are also provided. The invention further provides compositions, kits, and methods for their use to target sperm and modulate fertility. The invention further provides therapeutic compositions and methods for the use of *FSP95* polynucleotides for targeted heterologous gene expression.

The invention further encompasses the use of nucleotides encoding FSP95 proteins and peptides, as well as antibodies to FSP95 (which can, for example, act as agonists or antagonists), ligands that bind to FSP95 or modulate the function, activity or expression of FSP95. In addition, regulatory nucleotides and nucleotides encoding FSP95 polypeptides or one or more functional domains of FSP95 or fragments thereof, *e.g.*, the RII binding site, are effective in gene therapy, or for delivery of heterologous gene products to a cellular or subcellular locale.

In particular, embodiments of the invention described in the subsections below encompasses FSP95, polypeptides or peptides corresponding to functional domains of FSP95 (*e.g.*, a ligand-binding domain, such as the RII binding site), mutated, truncated or deleted (*e.g.* with one or more functional domains or portions thereof deleted), FSP95

fusion proteins, nucleotide sequences encoding such products, and host cell expression systems that can produce such FSP95 products.

The invention also encompasses antibodies, including anti-idiotypic antibodies, antagonists and agonists, as well as compounds or nucleotide constructs that inhibit expression of the FSP95 gene (transcription factor inhibitors, antisense and ribozyme molecules, or gene or regulatory sequence replacement constructs), or promote expression of FSP95 (*e.g.*, expression constructs in which FSP95 coding sequences are operatively associated with expression control elements such as promoters, promoter/enhancers, *etc.*). The invention also provides host cells and animals genetically engineered to express the human (or mutants thereof) FSP95 regulatory or protein coding sequences, or to inhibit or "knock-out" expression of the animal's endogenous FSP95.

The FSP95 products and fusion protein products, (*i.e.*, fusions of the proteins or a domain of the protein, *e.g.*, the RII binding site), antibodies and anti-idiotypic antibodies (including Fab fragments), modulators and ligands can be used as therapeutics to modulate fertility.

The present invention provides methods of screening for agents, small molecules, or proteins that interact with FSP95. The invention encompasses both *in vivo* and *in vitro* assays to screen small molecules, compounds, recombinant proteins, peptides, nucleic acids, antibodies *etc.* which bind to or modulate the activity of FSP95 and are thus useful as therapeutics or diagnostic markers for fertility.

The *FSP95* polynucleotides (*i.e.*, coding regions or regulatory regions of the *FSP95* gene) and fusion protein products, (*i.e.*, fusions of the FSP95 proteins or a domain of the protein, *e.g.*, the RII binding site, to another heterologous polypeptide), antibodies (including, but without limitation, anti-idiotypic antibodies, and Fab fragments), modulators and ligands can be used for drug delivery or gene therapy. Thus, the invention also encompasses pharmaceutical formulations and methods for contraception and treating infertility and cancer.

Various aspects of the invention are described in greater detail in the subsections below.

5.1 *FSP95* NUCLEIC ACIDS

The invention provides the nucleotide sequences of *FSP95* nucleic acids which were identified by screening a human testicular cDNA library screened with a degenerate primers designed from a peptide sequence obtained from purified FSP95. Nucleic acid sequences of the identified *FSP95* genes are described herein. As used herein, "a *FSP95* nucleic acid" refers to:

(a) a nucleic acid molecule containing the nucleotide sequence of *FSP95* shown in FIG. 3

(b) any nucleotide sequence that encodes a polypeptide containing the amino acid sequence of FSP95 shown in FIG.3 and FIG. 4;

(c) any nucleotide sequence that hybridizes to the complement of the DNA sequences that encode any of the amino acid sequences of FSP95 shown in FIG. 3 and FIG.4 under highly stringent conditions, *e.g.*, hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3); and/or

(d) any DNA sequence that hybridizes to the complement of the DNA sequences that encode any of the amino acid sequences in FSP95 shown in FIG. 3, under less stringent conditions, such as moderately stringent conditions, *e.g.*, washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*), and encodes a gene product functionally equivalent to a *FSP95* gene product.

The invention also includes nucleic acid molecules derived from mammalian nucleic acids, preferably DNA molecules, that hybridize to, and are therefore the complements of, the nucleotide sequences (a) through (d), in the preceding paragraph. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, *e.g.*, to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as *FSP95* nucleic acid antisense molecules, useful, for example, in *FSP95* gene regulation (for and/or as antisense primers in amplification reactions of *FSP95* nucleic acid sequences). With respect to *FSP95* gene regulation, such techniques can be used to regulate, for example, *a FSP95*-regulated pathway, in order to block cell proliferation associated with cancer. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for *FSP95* gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby, for example, the presence of a particular *FSP95* allele responsible for causing *a FSP95* related disorder, *e.g.*, fertility or proliferative disorders such as infertility or cancer, may be detected.

The nucleic acid molecules of the invention further include nucleotide sequences that encode polypeptides having at least 30%, 35%, 40%, 45%, 50%, 55%, 60%,

65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or higher amino acid sequence identity to the polypeptides encoded by the *FSP95* nucleotide sequences of (a)-(d) above.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical overlapping positions/total # of overlapping positions x 100%). In one embodiment, the two sequences are the same length.

The determination of percent identity between two sequences can also be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul, 1990, Proc. Natl. Acad. Sci. USA 87:2264-2268, modified as in Karlin and Altschul, 1993, Proc. Natl. Acad. Sci. USA 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al., 1990, J. Mol. Biol. 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., 1997, *Nucleic Acids Res.25*:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules (Altschul et al., 1997, *supra*). When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used (see http://www.ncbi.nlm.nih.gov). Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, 1988, *CABIOS 4*:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically only exact matches are counted.

The invention further includes fragments of any of the nucleotide sequences disclosed herein.

In a specific embodiment, the fragment of a *FSP95* nucleic acid encodes a *FSP95* PKA RII subunit binding domain sequence.

As used herein, a *FSP95* PKA RII subunit binding domain sequence includes:

(a) any DNA sequence that encodes a peptide, *e.g.*, a PKA RII subunit binding domain peptide, comprising amino acids 318-335 (LKKVLLKHAKEVVSDLID), 368-385 (QKATDIMDAMLRKLYNVM), or 671-688 (EHLMNSVMKLCVIIAKSC);

(B) any DNA sequence that hybridizes to the complement of the DNA sequences that encode the PKA RII subunit binding domain sequences described in (a) or (b) under highly stringent conditions, *e.g.*, hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, at p. 2.10.3); and/or

(c) any DNA sequence that hybridizes to the complement of the DNA sequences that encode the PKA RII subunit binding domain sequences described in (a) or (b) under less stringent conditions, such as moderately stringent conditions, *e.g.*, washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*), and encodes a gene product functionally equivalent to a *FSP95* PKA RII subunit binding domain.

The invention encompasses the *FSP95* PKA RII subunit binding domain sequences, in isolated or purified form, as well as compositions containing such PKA RII subunit binding domain sequences operatively associated with a nucleic acid encoding a protein or polypeptide heterologous to FSP95.

FSP95 sequences of the present invention are derived from a eukaryotic genome, preferably a mammalian genome, and more preferably a human genome. In a specific embodiment, the nucleotides of the present invention encompass any DNA sequence derived from a mammalian genome which hybridizes under highly stringent

conditions to the nucleotide sequence shown in FIG. 3. In a specific embodiment, the nucleotides of the present invention encompass any DNA sequence derived from a mammalian genome which hybridize under highly stringent conditions to the nucleotide sequence shown in FIG. 3 and encodes a gene product involved sperm motility, and contains a PKA RII subunit binding domain.

When referring to a nucleic acid which encodes a given amino acid sequence, it is to be understood that the nucleic acid need not only be a cDNA molecule, but can also, for example, refer to a gDNA sequence from which an mRNA species is transcribed that is processed to encode the given amino acid sequence.

The invention further includes regulator nucleic acids of the *FSP95* gene. The genomic sequence of the *FSP95* gene contains regulatory sequences in the non-coding 5'- flanking region. The 5'-regulatory sequences of the *FSP95* gene comprise the polynucleotide sequences located between the nucleotide in position -5000, -3000, -2000, -1000, or -500, and the nucleotide in position +10, +100, or +300, of the nucleotide sequence of the nucleotide sequence shown in FIG. 3.

The invention also encompasses:

(a) vectors that contain any of the foregoing *FSP95* coding sequences and/or their complements (*i.e.*, antisense);

(b) expression vectors that contain any of the foregoing *FSP95* coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and

(c) genetically engineered host cells that contain any of the foregoing *FSP95* coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell.

As used herein, regulatory elements include but are not limited to inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include but are not limited to the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage A, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast α-mating factors.

In one embodiment, the *FSP95* nucleic acid sequences of the invention are mammalian nucleic acid sequences, with human sequences being preferred.

In yet another embodiment, the *FSP95* nucleic acid sequences of the invention are nucleic acid sequences encoding FSP95 gene products containing polypeptide

portions corresponding to (that is, polypeptide portions exhibiting amino acid sequence similarity to) the amino acid sequences depicted in FIG. 3 and FIG. 4, wherein the corresponding portion exhibits greater than about 50% amino acid identity with the depicted sequence, averaged across the FSP95 gene product's entire length.

In specific embodiments, FSP95 encoding nucleic acids comprise the cDNA sequences of the nucleotide sequences shown in FIG. 3 or the coding regions thereof, or nucleic acids encoding a FSP95 protein (*e.g.*, a protein having the amino acid sequences depicted in FIG. 3 and FIG. 4). The invention provides isolated or purified nucleic acids consisting of at least 8 nucleotides (*i.e.*, a hybridizable portion) of a *FSP95* nucleic acid sequence; in other embodiments, the nucleic acids consist of at least 25 (continuous) nucleotides, 50, 100, 150, 200, 300, 400, 500, 600, 700, 800,900, 1000, 1500, 2000, 2500, 2900, or 2941 contiguous nucleotides of a *FSP95* sequence, or a full-length *FSP95* coding sequence. For example, in one embodiment the invention provides isolated or purified nucleic acids consisting of nucleotides 1-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-750, 751-1000, 1001-1501, or 1501-2000, 2001-2500, or 2501-2941, of the nucleotide sequence shown in FIG. 3. In another embodiment, the nucleic acids are smaller than 25, 50, 75, 100, 200, 300, 400, 500, 1000, 1500, 2000, 2500, or 2941 nucleotides in length. Nucleic acids can be single or double stranded. The invention also provides nucleic acids hybridizable to or complementary to the foregoing sequences. In specific aspects, nucleic acids are provided which comprise a sequence complementary to at least 10, 25, 50, 100, 200, 300, 400, 500, 1000, 1500, 2000, 2500 nucleotides or the entire coding region of a *FSP95* gene. As will be understood by those skilled in the art, the invention also encompasses those genomic DNA sequences which give rise to the cDNA sequences of the nucleotide sequences shown in FIG. 3 described above.

In addition to the human *FSP95* nucleic acid sequences disclosed in the nucleotide sequences shown in FIG. 3, other *FSP95* nucleic acid sequences can be identified and readily isolated, without undue experimentation, by molecular biological techniques well known in the art, used in conjunction with the *FSP95* nucleic acid sequences disclosed herein. These other sequences are encompassed by the present invention. For example, additional human *FSP95* nucleic acid sequences at the same or at different genetic loci as those disclosed in the nucleotide sequences shown in FIG. 3 can be isolated readily. There can exist, for example, genes at other genetic or physical loci within the human genome that encode proteins that have extensive homology to one or more domains of the *FSP95* gene products and that encode gene products functionally equivalent to a *FSP95* gene product. Further, homologous *FSP95* nucleic acid sequences present in other species can be identified and isolated readily.

With respect to identification and isolation of *FSP95* nucleic acid sequences present at the same genetic or physical locus as those sequences disclosed in the nucleotide sequences shown in FIG. 3, such sequences can, for example, be obtained readily by utilizing standard sequencing and bacterial artificial chromosome (BAC) and P1 artificial chromosome (PAC) technologies.

With respect to the cloning of a *FSP95* gene or nucleic acid homologue in human or other species (*e.g.*, mouse), the isolated *FSP95* nucleic acid sequences disclosed herein may be labeled and used to screen a cDNA library constructed from mRNA obtained from appropriate cells or tissues (*e.g.*, testes) derived from the organism (*e.g.*, mouse) of interest. The hybridization conditions used should be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived.

Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook, *et al.*, 1989, Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Press, N.Y.; and Ausubel, *et al.*, *supra*. Further, a *FSP95* gene homologue may be isolated from, for example, human nucleic acid, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within any *FSP95* gene product disclosed herein.

The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a *FSP95* gene nucleic acid sequence. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to isolate genomic clones via the screening of a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*, testis). A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C

primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies that may be used, see *e.g.*, Sambrook *et al.*, *supra*.

FSP95 nucleic acid sequences may additionally be used to identify mutant *FSP95* gene alleles. Such mutant alleles may be isolated from individuals either known or proposed to have a genotype that contributes to the symptoms of a *FSP95* gene disorder, such as fertility disorders, for example. Such alleles are encompassed by the present invention.

FSP95 alleles may be identified by single strand conformational polymorphism (SSCP) mutation detection techniques, Southern blot, and/or PCR amplification techniques. Primers can routinely be designed to amplify overlapping regions of the whole *FSP95* sequence including the promoter region. In one embodiment, primers are designed to cover the exon-intron boundaries such that, first, coding regions can be scanned for mutations. Genomic DNA isolated from lymphocytes of normal and affected individuals is used as PCR template. PCR products from normal and affected individuals are compared, either by single strand conformational polymorphism (SSCP) mutation detection techniques and/or by sequencing. SSCP analysis can be performed as follows: 100 ng of genomic DNA is amplified in a 10 μl reaction, adding 10 pmols of each primer, 0.5 U of Taq DNA polymerase (Promega), 1 μCi of α-[^{32}P]dCTP (NEN; specific activity, 3000 Ci/mmol), in 2.5 μM dNTPs (Pharmacia), 10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1 mM MgCl2, 0.01% gelatin, final concentration. Thirty cycles of denaturation (94°C), annealing (56°C to 64°C, depending on primer melting temperature), and extension (72°C) is carried out in a thermal-cycler (MJ Research, Boston, MA, USA), followed by a 7 min final extension at 72°C. Two microliters of the reaction mixture is diluted in 0.1% SDS, 10 mM EDTA and then mixed 1: 1 with a sequencing stop solution containing 20 mM NaOH. Samples are heated at 95°C for 5 min, chilled on ice for 3 min and then 3 μl will be loaded onto a 6% acrylamide/TBE gel containing 5% (v/v) glycerol. Gels are run at 8 W for 12-15 h at room temperature. Autoradiography is performed by exposure to film at -70°C with intensifying scree for different periods of time. The mutations responsible for the loss or alteration of function of the mutant *FSP95* gene product can then be ascertained.

Alternatively, a cDNA of a mutant *FSP95* gene may be isolated, for example, using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant *FSP95* allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a

suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant *FSP95* allele to that of the normal *FSP95* allele, the mutation(s) responsible for the loss or alteration of function of the mutant *FSP95* gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant *FSP95* allele, or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant *FSP95* allele. An unimpaired *FSP95* gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant *FSP95* allele in such libraries. Clones containing the mutant *FSP95* nucleic acid sequences may then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant *FSP95* allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal *FSP95* gene product, as described, below, in Section 5.3. (For screening techniques, see, for example, Harlow and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.)

Nucleic acids encoding derivatives and analogs of FSP95 proteins, and FSP95 antisense nucleic acids can be isolated by the methods recited above. As used herein, a "nucleic acid encoding a fragment or portion of a FSP95 protein" shall be construed as referring to a nucleic acid encoding only the recited fragment or portion of the FSP95 and not the other contiguous portions of the FSP95 protein as a continuous sequence.

Fragments of *FSP95* nucleic acids comprising regions conserved between (*i.e.*, with homology to) other *FSP95* nucleic acids, of the same or different species, are also provided. Nucleic acids encoding one or more FSP95 domains can be isolated by the methods recited above.

In cases where a *FSP95* mutation results in an expressed gene product with altered function (*e.g.*, as a result of a missense or a frameshift mutation), a polyclonal set of anti-*FSP95* gene product antibodies are likely to cross-react with the mutant *FSP95* gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

5.2 FSP95 PROTEINS AND POLYPEPTIDES

The invention further provides *FSP95*-encoded proteins and amino acid sequences, as well as derivatives (*e.g.*, fragments) and analogs thereof. In particular, the invention provides FSP95 derivatives and analogs which are functionally active, i.e., they are capable of displaying one or more functional activities associated with a full-length (wild-type) FSP95 protein. Such functional activities include, but are not limited to, binding to the RII subunit of PKA, phosphorylation, modulating sperm motility, sperm marker, antigenicity (ability to bind to an anti-FSP95 antibody or compete with FSP95 for binding), immunogenicity (ability to generate antibody which binds to FSP95). The invention further provides fragments (and derivatives and analogs thereof) of FSP95 which comprise one or more domains of a FSP95 protein. In a specific embodiment, the FSP95 protein is a human protein.

The amino acid sequences depicted in FIG. 3 and FIG.4 represent *FSP95* gene products. The *FSP95* gene product, sometimes referred to herein as a "FSP95", includes those products encoded by the *FSP95* nucleic acid sequences described in Section 5.1, above. In accordance with the present invention, the nucleic acid sequences encoding the *FSP95* gene products are derived from eukaryotic genomes, including mammalian genomes. In a preferred embodiment the nucleic acid sequences encoding the *FSP95* products are derived from the human genome.

FSP95 proteins, polypeptides and peptide fragments thereof, can be prepared for a variety of uses. For example, such molecules can be used for the generation of antibodies, for use in diagnostic and therapeutic assays, for the identification of other sperm gene products involved in sperm motility, or for the identification of compounds that modulate sperm motility.

In addition, *FSP95* products of the present invention may include proteins that represent functionally equivalent (see Section 5.1 for a definition) products. Functionally equivalent *FSP95* products may contain deletions, including internal deletions, additions, including additions yielding fusion proteins, or substitutions of amino acid residues within and/or adjacent to the amino acid sequence encoded by the *FSP95* nucleic acid sequences described, above, in Section 5.1, but that result in a "silent" change, in that the change produces a functionally equivalent *FSP95* gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include

glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Specific functional variants of FSP95 proteins encompassed by the present invention include immunoreactive proteins with the following molecular weights (MW) in kiloDaltons, and isoloelectric points (pI), obtainable from human sperm preparations: a protein

Alternatively, where alteration of function is desired, deletion or non-conservative alterations can be engineered to produce altered *FSP95* gene products. Such alterations can, for example, alter one or more of the biological functions of the *FSP95* product. Further, such alterations can be selected so as to generate *FSP95* products that are better suited for expression, scale up, etc. in the host cells chosen. For example, in one embodiment, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges. In another embodiment, tyrosine residues can be deleted or substituted with another amino acid residue in order to eliminate tyrosine phosphorylation.

Peptides and/or proteins corresponding to one or more domains of a FSP95 protein as well as fusion proteins in which a FSP95 protein or a portion of a FSP95 protein such as a truncated FSP95 protein or peptide or a FSP95 protein domain, is fused to an unrelated protein are also within the scope of this invention. Such proteins and peptides can be designed on the basis of the FSP95 nucleotide sequence disclosed in Section 5.1, above, and/or on the basis of the FSP95 amino acid sequence disclosed herein. Fusion proteins include, but are not limited to, IgFc fusions which stabilize the FSP95 protein or peptide and prolong half life *in vivo*; or fusions to any amino acid sequence that allows the fusion protein to be anchored to the cell membrane; or fusions of FSP95 protein domains to an enzyme, fluorescent protein, luminescent protein, or a flag epitope protein or peptide which provides a marker function.

FSP95 proteins of the invention also include FSP95 protein sequences wherein domains encoded by at least one exon of the cDNA sequence, or fragments thereof, have been deleted.

The FSP95 polypeptides of the invention can further comprise posttranslational modifications, including, but not limited to stearation, myristylations, palmitation, glycosylations, acetylations, and phosphorylations. In a preferred embodiment, FSP95 polypeptides are modified to make them membrane permeable using techniques such as stearation, myristylation, palmitation, and incorporating lipophilic amino acids. If the native FSP95 protein does not have recognition motifs that allow such modifications, it

would be routine for one skilled in the art to introduce into a FSP95 gene nucleotide sequences that encode motifs such as enzyme recognition signals so as to produce a modified FSP95 gene product.

5.3 PREPARATION OF PEPTIDES AND PEPTIDE ANALOGUES

5.3.1 CHEMICAL SYNTHESIS

The peptides of the invention or analogues thereof, may be prepared using virtually any art-known technique for the preparation of peptides and peptide analogues. For example, the peptides may be prepared in linear form using conventional solution or solid phase peptide syntheses and cleaved from the resin followed by purification procedures (Creighton, 1983, Protein Structures And Molecular Principles, W.H. Freeman and Co., N.Y.). Suitable procedures for synthesizing the peptides described herein are well known in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure and mass spectroscopy).

In addition, analogues and derivatives of the peptides can be chemically synthesized. The linkage between each amino acid of the peptides of the invention may be an amide, a substituted amide or an isostere of amide. Nonclassical amino acids or chemical amino acid analogues can be introduced as a substitution or addition into the sequence. Non-classical amino acids include, but are not limited to, the D-isomers of the common amino acids, α-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ-Abu, ∈-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β-methyl amino acids, Cα-methyl amino acids, Nα-methyl amino acids, and amino acid analogues in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

Cyclized peptides may be formed by the addition of Cys residues to the termini of linear peptides. Formation of disulfide linkages, if desired, is generally conducted in the presence of mild oxidizing agents. Chemical oxidizing agents may be used, or the compounds may simply be exposed to atmospheric oxygen to effect these linkages. Various methods are known in the art, including those described, for example, by Tam, J.P. *et al.*, 1979, *Synthesis* 955-957; Stewart *et al.*, 1984, *Solid Phase Peptide Synthesis*, 2d Ed., Pierce Chemical Company Rockford, IL; Ahmed *et al.*, 1975, *J. Biol. Chem.* 250:8477-8482; and Pennington *et al.*, 1991 *Peptides 1990* 164-166, Giralt and

Andreu, Eds., ESCOM Leiden, The Netherlands. An additional alternative is described by Kamber *et al.*, 1980, *Helv Chim Acta* 63:899-915. A method conducted on solid supports is described by Albericio, 1985, *Int. J. Peptide Protein Res.* 26:92-97. Any of these methods may be used to form disulfide linkages in the peptides of the invention.

5.3.2 RECOMBINANT SYNTHESIS

If the peptide is composed entirely of gene-encoded amino acids, or a portion of it is so composed, the peptide or the relevant portion may also be synthesized using conventional recombinant genetic engineering techniques.

For recombinant production, a polynucleotide sequence encoding a linear form of the peptide is inserted into an appropriate expression vehicle, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence, or in the case of an RNA viral vector, the necessary elements for replication and translation. The expression vehicle is then transfected into a suitable target cell which will express the peptide. Depending on the expression system used, the expressed peptide is then isolated by procedures well-established in the art. Methods for recombinant protein and peptide production are well known in the art (*see*, *e.g.*, Maniatis *et al.*, 1989, *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory, N.Y.; and Ausubel *et al.*, 1989, *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley Interscience, N.Y.). The coding sequence for human GAD65 has been described (Bu et al., 1992, Proc. Natl. Acad. Sci. U.S.A. 89:2115-2119; Bu and Tobin, 1994, Genomics 21:222-228). Methods for introducing codon substitutions to the native sequence in order to encode an antagonistic peptide based on the disclosure herein are well known to those skilled in the art.

5.3.3 RECOMBINANT FSP95 PROTEINS, VECTORS, AND CELLS

The FSP95 gene products, peptide fragments thereof and fusion proteins thereof, may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the FSP95 gene polypeptides, peptides, fusion peptide and fusion polypeptides of the invention by expressing nucleic acid containing *FSP95* gene sequences are described herein. Methods that are well known to those skilled in the art can be used to construct expression vectors containing FSP95 gene product coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook, et al., 1989, *supra*,

and Ausubel, *et al.*, 1989, *supra*. Alternatively, RNA capable of encoding *FSP95* gene product sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, ed., IRL Press, Oxford.

A variety of host-expression vector systems may be utilized to express the FSP95 gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells that may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the FSP95 gene product of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing FSP95 gene product coding sequences; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the FSP95 gene product coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the FSP95 gene product coding sequences; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing FSP95 gene product coding sequences; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the FSP95 gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of FSP95 protein or for raising antibodies to FSP95 protein, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, 1983, EMBO J. 2, 1791), in which the *FSP95* gene product coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, 1985, Nucleic Acids Res. 13, 3101-3109; Van Heeke and Schuster, 1989, J. Biol. Chem. 264, 5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The

pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica*, nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. *FSP95* gene coding sequences may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of *FSP95* gene coding sequences will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (*e.g.*, see Smith, *et al.*, 1983, J. Virol. 46, 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a *FSP95* gene coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing *FSP95* gene product in infected hosts (*e.g.*, see Logan and Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted *FSP95* gene product coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire *FSP95* gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the *FSP95* gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner, *et al.*, 1987, Methods in Enzymol. 153, 516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein

products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, and WI38.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the *FSP95* gene product may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci that in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the *FSP95* gene product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the *FSP95* gene product.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, 1977, Cell 11: 223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48: 2026), and adenine phosphoribosyltransferase (Lowy, *et al.*, 1980, Cell 22: 817) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to mycophenolic acid (Mulligan and Berg, 1981, Proc. Natl. Acad. Sci. USA 78: 2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, 1981, J. Mol. Biol. 150: 1); and hygro, which confers resistance to hygromycin (Santerre, *et al.*, 1984, Gene 30: 147).

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht, *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, *et al.*, 1991, Proc. Natl. Acad. Sci. USA 88,: 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination

plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

Alternatively, the expression characteristics of an endogenous *FSP95* gene within a cell, cell line, or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous *FSP95* gene. For example, an endogenous *FSP95* gene which is normally "transcriptionally silent", *i.e.*, a *FSP95* gene which is normally not expressed, or is expressed only at very low levels in a cell, cell line, or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell, cell line, or microorganism. Alternatively, a transcriptionally silent, endogenous *FSP95* gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous *FSP95* gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

FSP95 gene products can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, sheep, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate *FSP95* transgenic animals. The term "transgenic," as used herein, refers to animals expressing *FSP95* gene sequences from a different species (*e.g.*, mice expressing human *FSP95* sequences), as well as animals that have been genetically engineered to overexpress endogenous (*i.e.*, same species) *FSP95* sequences or animals that have been genetically engineered to no longer express endogenous *FSP95* gene sequences (*i.e.*, "knock-out" animals), and their progeny.

Any technique known in the art may be used to introduce an *FSP95* gene transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe and Wagner, 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (van der Putten, *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82: 6148-6152); gene targeting in embryonic stem cells (Thompson, *et al.*, 1989, Cell 56: 313-321); electroporation of embryos (Lo, 1983, Mol. Cell. Biol. 3: 1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell

57: 717-723). For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115: 171-229).

Any technique known in the art may be used to produce transgenic animal clones containing an *FSP95* transgene, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal or adult cells induced to quiescence (Campbell, *et al.*, 1996, Nature 380,: 64-66; Wilmut, *et al.*, Nature 385,: 810-813).

The present invention provides for transgenic animals that carry a *FSP95* transgene in all their cells, as well as animals that carry the transgene in some, but not all their cells, *i.e.*, mosaic animals. The transgene may be integrated as a single transgene or in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.* (Lasko, *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89, 6232-6236). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the *FSP95* gene transgene be integrated into the chromosomal site of the endogenous *FSP95* gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous *FSP95* gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous *FSP95* gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous *FSP95* gene in only that cell type, by following, for example, the teaching of Gu, *et al.* (Gu, *et al.*, 1994, Science 265, 103-106). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the phenotypic expression of the recombinant *FSP95* gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques that include but are not limited to Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR (reverse transcriptase PCR). Samples of *FSP95* gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the *FSP95* transgene product.

FSP95 gene products, or peptide fragments thereof, can be prepared for a variety of uses. For example, such gene products, or peptide fragments thereof, can be used

for the generation of antibodies, in diagnostic assays, or for mapping and the identification of other cellular or extracellular gene products involved in the sperm motility. Such *FSP95* gene products include but are not limited to soluble derivatives such as peptides or polypeptides corresponding to one or more domains of the *FSP95* gene product, particularly *FSP95* gene products, that are modified such that they are deleted for one or more hydrophobic domains. Alternatively, antibodies to the *FSP95* protein or anti-idiotypic antibodies that mimic the *FSP95* gene product (including Fab fragments), antagonists or agonists can be used to treat *FSP95*-related disorders, such as fertility disoders. In yet another approach, *FSP95* gene products can be directly administered to a subject to modulate sperm motility, for example, for use as, a contraceptive or to increase fertility. In another embodiment, nucleotide constructs encoding such *FSP95* gene products can be used to genetically engineer host cells to express such *FSP95* gene products *in vivo*; these genetically engineered cells can function as "bioreactors" in the body delivering a continuous supply of *FSP95* gene product, *FSP95* peptides, or soluble *FSP95* polypeptides.

5.3.4 COMPOSITIONS CONTAINING ONE OR MORE DOMAINS OF FSP95

In a specific embodiment, the invention provides FSP95 fragments or analogs and derivatives of such fragments, that comprise, or alternatively consist of, one or more domains of a FSP95 protein, for example, a nuclear translocation domain.

A specific embodiment provides molecules comprising specific fragments of FSP95 that are those fragments in the respective FSP95 protein most homologous to specific fragments of a human FSP95 protein. A fragment comprising a domain of a FSP95 homolog can be identified by protein analysis methods as described in Sections 5.3.2 or 6.

In a specific embodiment, the invention provides a fragment, derivative or analog of a FSP95 protein that has a functional RII subunit binding domain. See above Section 5.1 for nucleic acid sequences encoding a functional RII subunit binding domain. In another specific embodiment, the invention provides fusion proteins comprising an FSP95 RII subunit binding domain that has been operatively linked to a heterologous protein. The fusion proteins are useful to cause the heterologous protein to be translocated to the nucleus. It is known that a proline followed within the next six residues by three lysines, has been shown to direct translocation of proteins into the nucleus (Hicks, G. R., and Raikhel, N. V. (1995). Protein import into the nucleus: an integrated view. *Annu. Rev. Cell. Dev. Biol.* 11, 155-158). See *infra* Section 5.6 for a more detailed discussion.

In another specific embodiment, a molecule is provided that comprises one or more domains (or functional portion thereof) of a FSP95 protein but that also lacks one

or more domains (or functional portion thereof) of a FSP95 protein. In particular examples, FSP95 protein derivatives are provided that lack an RII subunit binding domain.

5.4 FSP95 ANTIBODIES

According to the invention, FSP95, its fragments or other derivatives, or analogs thereof, may be used as an immunogen to generate antibodies which immunospecifically bind such an immunogen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. In a specific embodiment, antibodies to a human FSP95 protein are produced. In another embodiment, antibodies to a domain (*e.g.*, the RII subunit binding domain) of a FSP95 are produced.

Various procedures known in the art may be used for the production of polyclonal antibodies to a FSP95 or derivative or analog. In a particular embodiment, rabbit polyclonal antibodies to an epitope of an FSP95 encoded by a sequence as shown in FIG.3 or a subsequence thereof, can be obtained. For the production of antibody, various host animals can be immunized by injection with the native FSP95, or a synthetic version, or derivative (*e.g.*, fragment) thereof, including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response, depending on the host species, and including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. For preparation of monoclonal antibodies directed toward an FSP95 sequence or analog thereof, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals (see PCT International Publication No. WO 89/12690, published December 12, 1989). According to the invention, human antibodies may be used and can be obtained by using human hybridomas (Cote et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030) or by transforming human B cells with EBV virus *in vitro* (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, pp. 77-96). In fact, according to the invention, techniques developed for the

production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851-6855; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule specific for FSP95 together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention.

According to the invention, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce FSP95-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse et al., 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for FSP95, or derivatives, or analogs.

Antibody fragments which contain the idiotype of the molecule can be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragment, the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent, and Fv fragments.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.* ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognize a specific domain of an FSP95, one may assay generated hybridomas for a product which binds to an FSP95 fragment containing such domain. For selection of an antibody that specifically binds a first FSP95 homolog but which does not specifically bind a different FSP95 homolog, one can select on the basis of positive binding to the first FSP95 homolog and a lack of binding to the second FSP95 homolog.

Antibodies specific to a domain of an FSP95 are also provided, such as a PKA RII subunit binding domain.

The foregoing antibodies can be used in methods known in the art relating to the identification of sperm containing FSP95 protein, separation of sperm, and the localization and activity of the FSP95 polypeptides of the invention, *e.g.*, for imaging these proteins, in diagnostic methods, measuring levels thereof in appropriate physiological samples etc.

5.5 FSP95 LIGANDS

Any of a variety of exogenous compounds, both naturally occurring and/or synthetic (*e.g.*, libraries of small molecules or peptides, may be screened for FSP95 binding capacity.

All of these methods comprise the step of mixing an FSP95 protein or fragment with test compounds, allowing time for any binding to occur, and assaying for any bound complexes. All such methods are enabled by the present disclosure of substantially pure FSP95 proteins, substantially pure functional domain fragments, fusion proteins, antibodies, and methods of making and using the same.

In one embodiment, peptide libraries may be used to screen for agonists or antagonists of the FSP95 of the present invention diversity libraries, such as random or combinatorial peptide or nonpeptide libraries can be screened for molecules that specifically bind to FSP95. Many libraries are known in the art that can be used, *e.g.*, chemically synthesized libraries, recombinant (*e.g.*, phage display libraries), and *in vitro* translation-based libraries.

Examples of chemically synthesized libraries are described in Fodor et al., 1991, Science 251:767-773; Houghten et al., 1991, Nature 354:84-86; Lam et al., 1991, Nature 354:82-84; Medynski, 1994, Bio/Technology 12:709-710; Gallop et al., 1994, J. Medicinal Chemistry 37(9):1233-1251; Ohlmeyer et al., 1993, Proc. Natl. Acad. Sci. USA 90:10922-10926; Erb et al., 1994, Proc. Natl. Acad. Sci. USA 91:11422-11426; Houghten et al., 1992, Biotechniques 13:412; Jayawickreme et al., 1994, Proc. Natl. Acad. Sci. USA 91:1614-1618; Salmon et al., 1993, Proc. Natl. Acad. Sci. USA 90:11708-11712; PCT Publication No. WO 93/20242; and Brenner and Lerner, 1992, Proc. Natl. Acad. Sci. USA 89:5381-5383.

Examples of phage display libraries are described in Scott & Smith, 1990, Science 249:386-390; Devlin et al., 1990, Science, 249:404-406; Christian, et al., 1992, J. Mol. Biol. 227:711-718; Lenstra, 1992, J. Immunol. Meth. 152:149-157; Kay et al., 1993, Gene 128:59-65; and PCT Publication No. WO 94/18318 dated August 18, 1994.

In vitro translation-based libraries include but are not limited to those described in PCT Publication No. WO 91/05058 dated April 18, 1991; and Mattheakis et al., 1994, Proc. Natl. Acad. Sci. USA 91:9022-9026.

By way of examples of nonpeptide libraries, a benzodiazepine library (*see e.g.*, Bunin et al., 1994, Proc. Natl. Acad. Sci. USA 91:4708-4712) can be adapted for use. Peptoid libraries (Simon et al., 1992, Proc. Natl. Acad. Sci. USA 89:9367-9371) can also be used. Another example of a library that can be used, in which the amide functionalities in

peptides have been permethylated to generate a chemically transformed combinatorial library, is described by Ostresh et al. (1994, Proc. Natl. Acad. Sci. USA 91:11138-11142).

Screening the libraries can be accomplished by any of a variety of commonly known methods. See, *e.g.*, the following references, which disclose screening of peptide libraries: Parmley & Smith, 1989, Adv. Exp. Med. Biol. 251:215-218; Scott & Smith, 1990, Science 249:386-390; Fowlkes et al., 1992; BioTechniques 13:422-427; Oldenburg et al., 1992, Proc. Natl. Acad. Sci. USA 89:5393-5397; Yu et al., 1994, Cell 76:933-945; Staudt et al., 1988, Science 241:577-580; Bock et al., 1992, Nature 355:564-566; Tuerk et al., 1992, Proc. Natl. Acad. Sci. USA 89:6988-6992; Ellington et al., 1992, Nature 355:850-852; US 5,096,815, US 5,223,409, and US 5,198,346, all to Ladner et al.; Rebar & Pabo, 1993, Science 263:671-673; and PCT Publication No. WO 94/18318.

In a specific embodiment, screening can be carried out by contacting the library members with a FSP95 protein (or nucleic acid or derivative) immobilized on a solid phase and harvesting those library members that bind to the protein (or nucleic acid or derivative). Examples of such screening methods, termed "panning" techniques are described by way of example in Parmley & Smith, 1988, Gene 73:305-318; Fowlkes et al., 1992, BioTechniques 13:422-427; PCT Publication No. WO 94/18318; and in references cited hereinabove.

In another embodiment, the two-hybrid system for selecting interacting proteins or peptides in yeast (Fields & Song, 1989, Nature 340:245-246; Chien et al., 1991, Proc. Natl. Acad. Sci. USA 88:9578-9582) can be used to identify molecules that specifically bind to an FSP95 protein or derivative.

Binding interactions between two or more components can be measured in a variety of ways. One approach is to label one of the components with an easily detectable label, place it together with the other component(s) in conditions under which they would normally interact, perform a separation step which separates bound labeled component from unbound labeled component, and then measure the amount of bound component. The effect of a test agent included in the binding reaction can be determined by comparing the amount of labeled component which binds in the presence of this agent to the amount which binds in its absence.

The separation step in this type of procedure can be accomplished in various ways. In one approach, (one of) the binding partner(s) for the labeled component can be immobilized on a solid phase prior to the binding reaction, and unbound labeled component can be removed after the binding reaction by washing the solid phase. Attachment of the binding partner to the solid phase can be accomplished in various ways known to those skilled in the art, including, but not limited to, chemical cross-linking, non-specific

adhesion to a plastic surface, interaction with an antibody attached to the solid phase, interaction between a ligand attached to the binding partner (such as biotin) and a ligand-binding protein (such as avidin or streptavidin) attached to the solid phase, and so on.

Alternatively, the separation step can be accomplished after the labeled component had been allowed to interact with its binding partner(s) in solution. If the size differences between the labeled component and its binding partner(s) permit such a separation, the separation can be achieved by passing the products of the binding reaction through an ultrafilter whose pores allow passage of unbound labeled component but not of its binding partner(s) or of labeled component bound to its partner(s). Separation can also be achieved using any reagent capable of capturing a binding partner of the labeled component from solution, such as an antibody against the binding partner, a ligand-binding protein which can interact with a ligand previously attached to the binding partner, and so on.

5.6 METHODS AND COMPOSITIONS FOR DIAGNOSTIC AND THERAPEUTIC USE OF FSP95 NUCLEIC ACIDS, PROTEINS, DERIVATIVES, ANTIBODIES AND MODULATORS

5.6.1 THERAPEUTIC USE OF FSP95 REGULATORY SEQUENCES FOR TARGETING GENES TO THE TESTIS

FSP95 promoter sequences can be used advantageously to drive spermatid-specific expression of heterologous gene products. A vector comprising the *FSP95* promoter nucleotide sequences operably linked to a heterologous gene can be useful for gene therapy and contraception. In one embodiment, these sequences can be used for contraceptive or sterilization purposes. *FSP95* promoter sequences can be inserted into a vector operatively linked to a gene that will kill the cell in which it is expressed. Examples of such genes are known in the art, including, but not limited to, spermicides and toxins. In another embodiment, such a vector can be used to target cancer cells of a patient with a testes-specific proliferative disorder or cancer to inhibit growth or kill the cancer cell.

In one embodiment, *FSP95* promoter sequences can be used to drive spermatid-specific expression of drugs or toxins using gene therapy techniques in cells of a patient with a testes-specific proliferative disorder or cancer to inhibit growth or kill the cancer cell. In another embodiment, gene therapy techniques using promoter constructs to drive spermatid-specific expression of drugs or toxins can be used for sterilization or contraception in the testis.

The genomic sequence of the *FSP95* gene contains regulatory sequences both in the non-coding 5′-flanking gene of polynucleotide sequence of FIG. 3 can be assessed by any known method. In one embodiment, methods for *FSP95* gene comprise the polynucleotide sequences located between the nucleotide in position -2000 and the nucleotide in position +100 of the nucleotide sequence of FIG. 3 or more preferably between positions -3000 and +200 of FIG. 3.

Methods for identifying the 5′-regulatory sequences of the *FSP95* polynucleotide fragments of FIG. 3 involved in the regulation of the expression of the *FSP95* gene are well-known to those skilled in the art (see Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). An example of a typical method that can be used involves a recombinant vector carrying a reporter gene and genomic sequences from the *FSP95* genomic promoter sequence of FIG. 3. Briefly, the expression of the reporter gene (for example, green fluorescent protein, luciferase, β-galactosidase, or chloramphenicol acetyl transferase) is detected when placed under the control of a biologically active polynucleotide fragment. Genomic sequences located upstream of the first exon of the gene may be cloned into any suitable promoter reporter vector, such as the pSEAPBasic, pSEAP-Enhancer, pβgal-Basic, pβgal-Enhancer, or pEGFP-1 Promoter Reporter vectors available from Clontech, or pGL2-basic or pGL3-basic promoterless luciferase reporter gene vector from Promega. Each of these promoter reporter vectors include multiple cloning sites positioned upstream of a reporter gene encoding a readily assayable protein such as secreted alkaline phosphatase, green fluorescent protein, luciferase, or β-galactosidase. The sequences upstream of the first *FSP95* exon are inserted into the cloning sites upstream of the reporter gene in both orientations and introduced into an appropriate host cell. The level of reporter protein is assayed and compared to the level obtained with a vector lacking an insert in the cloning site. The presence of an elevated expression level in the vector containing the insert with respect to the control vector indicates the presence of a promoter in the insert.

Promoter sequences within the 5' non-coding regions of the *FSP95* gene may be further defined by constructing nested 5' and/or 3' deletions using conventional techniques such as Exonuclease III or appropriate restriction endonuclease digestion. The resulting deletion fragments can be inserted into the promoter reporter vector to determine whether the deletion has reduced or obliterated promoter activity, such as described, for example, by Coles et al. (Hum. Mol. Genet., 7:791-800, 1998). In this way, the boundaries of the promoters may be defined. If desired, potential individual regulatory sites within the promoter may be identified using site directed mutagenesis or linker scanning to obliterate potential transcription factor binding sites within the promoter individually or in

combination. The effects of these mutations on transcription levels may be determined by inserting the mutations into cloning sites in promoter reporter vectors. This type of assays are well known to those skilled in the art (WO 97/17359, US 5,374,544, EP 582 796, US 5,698,389, US 5,643,746, US5,502,176, and US 5,266,488).

The activity and the specificity of the promoter of the *FSP95* gene can further be assessed by monitoring the expression level of a detectable polynucleotide operably linked to the FSP95 promoter in different types of cells and tissues. The detectable polynucleotide may be either a polynucleotide that specifically hybridizes with a predefined oligonucleotide probe, or a polynucleotide encoding a detectable protein, including a *FSP95* polypeptide or a fragment or a variant thereof. This type of assay is well known to those skilled in the art (US 5,502,176 and US 5,266,488).

Polynucleotides carrying the regulatory elements located both at the 5' end and at the 3' end of the *FSP95* gene coding region may be advantageously used to control the transcriptional and translational activity of an heterologous polynucleotide of interest, said polynucleotide being heterologous as regards to the *FSP95* regulatory region.

Thus, the present invention also provides a purified, isolated, and recombinant nucleic acid comprising a polynucleotide sequence located between the nucleotide in position -2000 and the nucleotide in position +100 of the nucleotide sequence of FIG. 3, or a sequence complementary thereto or a functionally active fragment thereof.

By a "functionally active" fragment of the sequence of FIG. 3 according to the present invention is intended a polynucleotide comprising or alternatively consisting of a fragment of said polynucleotide which is functional as a regulatory region for expressing a recombinant polypeptide or a recombinant polynucleotide in a recombinant cell host. For the purpose of the invention, a nucleic acid or polynucleotide is "functional" as a regulatory region for expressing a recombinant polypeptide or a recombinant polynucleotide if said regulatory polynucleotide contains nucleotide sequences which contain transcriptional and translational regulatory information, and such sequences are "operably linked" to nucleotide sequences which encode the desired polypeptide or the desired polynucleotide.

The regulatory polynucleotides according to the invention may be advantageously part of a recombinant expression vector that may be used to express a coding sequence in a desired host cell or host organism.

5.6.2 GENE REPLACEMENT DELIVERY OR THERAPY

With respect to testes-specific gene expression, *FSP95* gene regulatory sequences, described, above, in Section 5.6.1 can, for example, be utilized for the treatment of proliferative disorders such as testicular cancer. Such treatment can be administered, for

example, in the form of gene replacement therapy. Specifically, one or more copies of a normal *FSP95* gene or a portion of the *FSP95* gene that directs the production of a heterologous gene product that is toxic to the cell, may be inserted into the appropriate cells within a patient, using vectors that include, but are not limited to, adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Such gene replacement therapy techniques should be capable delivering *FSP95* gene sequences to testis-specific cell types within patients.

In another embodiment, techniques for delivery involve direct administration of such *FSP95* gene sequences to the site of the cells in which the *FSP95* nucleic acid sequences are to be expressed.

Alternatively, cells, preferably autologous cells, can be engineered to express *FSP95* nucleic acid sequences, and may then be introduced into a patient in positions appropriate for the amelioration of a testes-specific disorder, such as proliferative or differentiative disorders, *e.g.*, cancer and tumorigenesis. The expression of the heterologous gene sequences is controlled by the appropriate *FSP95* gene regulatory sequences to allow such expression in testes cells. When the cells to be administered are non-autologous cells, they can be administered using well known techniques that prevent a host immune response against the introduced cells from developing. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

As described in more detail in Section 6 below, motif analysis identified three PKA RII subunit binding domains at amino acids 318-335 (LKKVLLKHAKEVVSDLID), 368-385 (QKATDIMDAMLRKLYNVM), and 671-688 (EHLMNSVMKLCVIIAKSC) of FIG.3.

As indicated above in Sections 5.1 and 5.2, the present invention encompasses the nucleic acid sequences encoding the *FSP95* PKA RII subunit binding domains sequences, FSP95 PKA RII subunit binding domain polypeptides, and methods for using the PKA RII subunit binding domain peptides to inhibit sperm motility for use in contraceptive therapy.

Thus, the invention also encompasses isolated nucleic acid molecules comprising the DNA sequence of the PKA RII subunit binding domains described above in operative association with a nucleic acid encoding heterologous polynucleotide. The invention also encompasses isolated fusion proteins comprising the FSP95 consensus PKA RII subunit binding domains operatively associated with a heterologous polypeptide.

With respect to testes-specific gene therapy, the FSP95 PKA RII subunit binding domain sequences of the invention can be used, for example, in gene targeting heterologous sequences to the sperm. The invention further encompasses the use of the promoter sequences described above in gene targeting heterologous sequences specifically to the sperm fibrous sheath.

In one embodiment, the invention also encompasses the use of FSP95 promoter sequences either alone, respectively, in which the FSP95 promoter sequences can be used to drive spermatid-specific expression of drugs or toxins using gene therapy techniques in cells of a patient with a testes-specific proliferative disorder or cancer to inhibit growth or kill the cancer cell. In another embodiment, gene therapy techniques using promoter constructs either alone, or in combination with the nucleic acid sequences encoding the PKA RII subunit binding domain set forth in amino acids 318-335 (LKKVLLKHAKEVVSDLID), 368-385 (QKATDIMDAMLRKLYNVM), and 671-688 (EHLMNSVMKLCVIIAKSC) of FIG. 3, respectively, can be used to drive spermatid-specific expression of drugs or toxins can be used for sterilization or contraception in the testis.

5.6.3 TARGET DISORDERS

With respect to specific diseases and disorders for gene therapy diseases that can be treated or prevented by the methods of the present invention include, but are not limited to: diseases and disorders involving a fertility or infertility, deficiency in cell proliferation or in which cell proliferation is desired for treatment or prevention, and that can be treated or prevented by introduction of a heterologous gene in a testes-specific manner, include, but are not limited to, degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds; for example, to promote wound healing, or to promote regeneration in degenerated, lesioned or injured tissues, etc. In a specific embodiment, testicular disorders are treated. Other disorders that are contemplated within the scope of the invention are fertility disorders. Gene therapy for contraceptive or sterilization of otherwise normal patients or subjects for veterinary purposes are also within the scope of the invention. The subject is preferably an animal, including, but not limited to, animals such as foxes, rabbits, rodents, cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human. In a specific embodiment, a non-human mammal is the subject.

5.6.4 PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

Formulations and methods of administration that can be employed when the therapeutic comprises a nucleic acid are described in Sections 5.1-5.6 above; additional appropriate formulations and routes of administration can be selected from among those described hereinbelow.

Various delivery systems are known and can be used to administer a therapeutic of the invention, *e.g.*, encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the therapeutic, receptor-mediated endocytosis (see, *e.g.*, Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), construction of a therapeutic nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compositions of the invention into sperm cells by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, *e.g.*, by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, *e.g.*, in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection at the site (or former site) of a malignant tumor or neoplastic or pre-neoplastic tissue.

In another embodiment, the therapeutic can be delivered in a vesicle, in particular a liposome (*see* Langer, Science 249:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; *see* generally *ibid.*)

In yet another embodiment, the therapeutic can be delivered in a controlled release system. In one embodiment, a pump may be used (*see* Langer, *supra*; Sefton, CRC

Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)). In another embodiment, polymeric materials can be used (*see* Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J. Macromol. Sci. Rev. Macromol. Chem. 23:61 (1983); *see also* Levy et al., Science 228:190 (1985); During et al., Ann. Neurol. 25:351 (1989); Howard et al., J. Neurosurg. 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, *i.e.*, the testes, thus requiring only a fraction of the systemic dose (*see, e.g.*, Goodson, 1984, in Medical Applications of Controlled Release, *supra*, vol. 2, pp. 115-138).

Other controlled release systems are discussed in the review by Langer (1990, Science 249:1527-1533).

In a specific embodiment where the therapeutic is a nucleic acid encoding a protein therapeutic, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by use of a retroviral vector (see US 4,980,286), or by direct injection, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see *e.g.*, Joliot et al., 1991, Proc. Natl. Acad. Sci. USA 88:1864-1868), etc. Alternatively, a nucleic acid therapeutic can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a therapeutic, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica

gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the therapeutic, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The therapeutics of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and

the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

6. EXAMPLE

The following describes the cloning and sequencing of a novel fibrous sheath protein of 95 kDa (FSP95) which undergoes tyrosine phosphorylation during capacitation of human spermatozoa and has similarity to sperm A-kinase anchor proteins (AKAPs). FSP95 is both auto- and iso-antigenic in humans as it is recognized by sera containing antisperm antibodies from infertile men and women. The 853 amino acid residue protein has a calculated molecular weight of 94.6 kDa and a pI of 6.0 and contains multiple potential phosphorylation sites for protein kinase C and casein kinase II as well as one tyrosine kinase phosphorylation site at amino acid 435. Northern analysis detected a single transcript of ~ 3.0 kb, and Northern dot blot analysis of 50 human tissues, including 7 fetal tissues, revealed FSP95 mRNA expression only in testis. Employing sperm immobilization, indirect immunofluorescence and immunoelectronmicroscopy with antisera to purified recombinant FSP95, the protein was localized to the ribs of the fibrous sheath in the principal piece of sperm tail. The molecule showed clear evidence of generating more acidic isoforms during capacitation.

6.1 EXPERIMENTAL PROCEDURES

6.1.1 Preparation of spermatozoa, capacitation and extraction of sperm proteins:

Human semen samples were obtained from healthy volunteers by masturbation following 3 to 4 days of sexual abstinence. All samples tested negative for HIV. Ejaculates with normal semen volume, sperm count and motility were used in this

study (*World Health Organization (WHO) Laboratory Manual for the Examination of Human Semen and Cervical Mucus Interaction*, 3rd Ed., University Press, Cambridge, 1992). Following liquefaction of semen samples at room temperature, fresh sperm were separated from seminal plasma, germ cells, white blood cells and epithelial cells by Percoll (Pharmacia Biotech, Sweden) density gradient centrifugation and washed in Hams F-b medium (Gibco BRL, Life Technologies, NY) within 2 to 3 hours of ejaculation using procedures described in Naaby-Hansen et al., 1997, Biol. Reprod. 56:771-787, which is incorporated by reference herein in its entirety.

To obtain capacitated spermatozoa, motile sperm were prepared by the swim-up method for 60 min in BWW medium (Irvine Scientific, CA) containing no human serum albumin (Yanagimachi et al., 1979, Fertil. Steril. 31:562-574). Motile spermatozoa were subsequently collected from the supernatant and incubated in BWW medium supplemented with 3.0% human serum albumin (HSA) (Sigma, St. Louis, MO) at 37°C in 5% $C0_2$ for 6 hour to induce capacitation (Luconi et al, 1996, Biol. Reprod. 55:207-216, Fenichel et al., 1996, Biol. Reprod. 54:1405-1411, Ohashi et al., 1995, Fertil. Steril. 63:625-630). Capacitation was also performed in presence of 200 μM genistein, a protein tyrosine kinase inhibitor (Carrera et al., 1996, Dev. Biol. 180:284-296).

Fresh and capacitated spermatozoa were solubilized in a lysis buffer containing 9.8 M urea, 2% octyl-β-D-glucopyranoside (OBG) (ESA Inc, MA), 2% (v:v) ampholines, 100 mM dithiothreitol (Bio Rad, CA), 5 mM iodoacetamide (Sigma, MO), 5 mM EDTA, and four protease inhibitors - 2 mM PMSF (Sigma, St. Louis, MO), 3 mg/ml TLCK (Boehringer Mannheim, IN), 1.46 mM pepstatin A (Sigma), and 2.1 mM leupeptin (Sigma). Five hundred million sperm per milliliter were solubilized by shaking at 4°C for 45 min. Insoluble molecules were removed by centrifugation at 10,000 x g for 2 min and the supernatant was used for 2-D gel separation of sperm proteins according to known protocols (Naaby-Hansen et al., 1997, Biol. Reprod. 56:771-787).

6.1.2 Detection of tyrosine phosphorylated proteins

After a 6.0 hour incubation period in capacitation medium, sperm were solubilized in the lysis buffer (Naaby-Hansen, 1997, Biol. Reprod. 56:771-787). The solubilized proteins were separated by 2-D SDS-PAGE and stained with silver (Hochstrasser et al., 1988, Anal. Biochem. 173:424-435). The proteins were electrotransferred to nitrocellulose membranes (Towbin et al., 1979, Proc. Natl. Acad. Sci. USA 76:4350-4354) and the nonspecific binding sites on the membrane were blocked with 1% bovine serum albumin (BSA) in 10 mM Tris (pH 7.5), 0.1 M NaCl and 0.05% Tween 20 for 20 min at 37°C. The blocked membrane was then incubated with horseradish

peroxidase conjugated anti-phosphotyrosine monoclonal antibody RC-20 (Transduction Laboratories, KY) at 1:2500 dilution in the above buffer for 20 min at 37°C. The membrane was washed thoroughly and positive immunoreactive spots were detected by enhanced chemiluminescence assay (Amersham Corp., UK) according to the manufacturer's instructions. RC-20 is a well characterized monoclonal antibody with a specificity for phosphotyrosine previously documented (Ruff-Jamisson et al., 1993, Science 261:1733-1736, Burks et al., 1995, Science 269:83-86).

6.1.3 Immunoblotting of sperm proteins with infertile sera prescreened

Sera were obtained from infertile men and women with unexplained infertility having no diagnosed hormonal, infective or physical causes of their infertility. ASA in the infertile patient sera were detected by the indirect immunobead binding test (Patrizio et al., 1992, Fertil. Steril. 57:183-186). Sera chosen for immunoblotting analysis had a high immunobead binding test score i.e., more than 60% of spermatozoa revealed bead binding, indicating IgG, IgA and/or IgM specific ASA. The selected sera all had antibodies directed against the sperm head or the entire cell.

For immunoblotting, sperm proteins were extracted and separated by a 2-D SDS-PAGE system as described earlier (Naaby-Hansen et al., 1997, Biol. Reprod. 56:771-787). Following electrophoretic transfer of the proteins, the membranes were rinsed in phosphate buffer saline (PBS) (pH 7.4) and blocked with 5% dry milk in PBS-Tween (10 mM PBS with 0.05% Tween 20). The blots were then incubated with test serum diluted 1:1000 at 4°C overnight. A horseradish peroxidase conjugated goat antihuman IgG/IgM secondary antibody (Jackson ImniunoResearch Lab, PA) was then incubated with the blots for 1 h at a 1:5000 dilution in PBS-Tween and the immunoreactive spots were visualized by enhanced chemiluminescence using the manufacturer's protocol (Amersham Corp., UK).

6.1.4 Microsequencing of a 95 kDa tyrosine phosphorylated auto- and iso-antigen

Following identification of a 95 kDa phosphotyrosine containing auto- and iso-antigen on 2-D immunoblots, Coomassie stained protein spot was cored from a 1.5 mm thick 2-D SDS-PAGE gel. The gel core was diced into small fragments. The protein in these fragments was destained in methanol, reduced in 10 mM dithiothreitol and alkylated in 50 mM iodoacetamide in 0.1 M ammonium bicarbonate. After removing the reagents, the gel pieces were incubated with 12.5 ng/ p1 trypsin in 50 mM ammonium bicarbonate overnight at 37°C. Peptides were extracted from the gel pieces in 50% acetonitrile in 5% formic acid and microsequenced by tandem mass spectrometry by methods known in the art.

6.1.5 Cloning, sequencing and analysis of the cDNA

A completely degenerate deoxyinosine containing sense primer (5'-A/T-C/G-I GTI TT-C/T TT-C/T AA-C/T TT-C/T A/T/C-TI A/C-GI-3') was designed from one of the microsequences obtained by mass spectrometry, peptide number 6 (SVFFNFI/LR), and the oligonucleotide was synthesized by GIBCO BRL (Life Technologies, CA). Using this forward primer and an adapter primer, a 3' rapid amplification of cDNA ends (RACE) polymerase chain reaction (PCR) was performed using 0.25 ng of human testicular Marathon ready cDNA (Clontech, CA) in a 25 µl assay system for 40 cycles. Thermal cycling was done in a MJ Research (Watertown, MA) thermal cycler (PTC-200 DNA engine) using a program of one cycle at 94°C for 1.5 min and 40 cycles of 94°C for 30 sec, 46°C for 1 min and 68°C for 2 min. PCR products were separated on a 1.7% NuSieve (FMC, ME) agarose gel. A 1.0 kb DNA fragment was isolated, reamplified, cloned into - the pCR 2.1 -TOPO vector (Invitrogen, CA), and sequenced on a Perkin-Elmer Applied Biosystems DNA sequencer using big fluorescence dye terminator chemistry with Tao DNA polymerase. The 3' clone contained a 786 bp open reading frame and a 218 bp untranslated region. The 5' end of the cDNA was also amplified by 5' RACE PCR from the same template using an adapter primer and an antisense 3' gene specific primer (5'-AGC CTG GGG GGA GAA GAG GCC AAC GOT C-3') which was 118 bp downstream from the 5' end of the 1.0 kb 3' clone. A product of 2081 bp was obtained and cloned into the pCR 2.1-TOPO vector. The 5' clone revealed a 162 bp untranslated region and an open reading frame of 1919 bp. cDNA clones were sequenced in both directions using vector-derived and insert specific primers. The nucleotide and amino acid sequence data were analyzed using the Genetics Computer Group program package (*Wisconsin Package Program Manual,* Version 9, Madison, WI).

6.1.6 Northern and dot blot analysis

A Northern blot containing 2 µg of poly(A)+ RNA from eight selected human tissues and a normalized RNA dot blot containing 89 to 514 ng of mRNA from 50 different human tissues were obtained from Clontech. The Northern blot was probed with a ^{32}P labeled 2337 bp DNA fragment generated from the 5' end of the coding region and a 1007 bp DNA fragment from the 3' end of the cDNA. Probes were prepared by random oligonucleotide prime labeling (Feinberg and Vogelstein, 1983, Anal. Biochem. 132:6-13). Hybridization was performed in ExpressHyb solution (Clontech) at 68°C for 1 h followed by three washes in 2x SSC, 0.05% SDS at room temperature and two washes in 0.lx SSC, 0.1% SDS for 20 min at 50°C. The blot was exposed to X-ray film (Kodak, Blue XB-l) at -70°C for 60 h with two intensifying screens. The dot blot was probed with a ^{32}P labeled

2337 bp 5' cDNA. The blot was hybridized in ExpressHyb solution containing salmon sperm DNA and human placental Cot- 1 DNA overnight at 65°C. The blot was then washed three times in 2x SSC, 1% SDS at 65°C followed by two additional washes in 0.lx SSC, 0.5% SDS at 55°C. Autoradiography was performed by exposing the filter to X-ray film for 18 h at -70°C with two intensifying screens.

6.1.7 Expression of FSP95 in E. coli and purification of the recombinant protein

To express the molecule, 91% of the FSP95 cDNA (ORF) encoding amino acids from 1 to 779 was amplified by PCR from human testicular Marathon ready cDNA (Clontech, CA). Primers were designed to create an NdeI site at the 5' end and an XhoI site at the 3' end of the PCR product. The amplified DNA was ligated into the NdeI-XhoI sites of the pET28b expression vector (Novagen, WI). The resulting construct appended 28 amino acids from the vector including six residues of histidine tag on either side of the protein. The FSP95 coding sequence was preceded by the promotor for phage T7 RNA polymerase, an initiator ATG and a six consecutive His codons. The recombinant vector was introduced into the Escherichia coli strain NovaBlue(DE3) cells, which contains a chromosomal copy of T7 RNA polymerase under the control of the lac promotor. The expression plasmid was sequenced at the 5' and 3' end to verify the reading frame of the construct.

A single positive colony was used to inoculate 10 liters of LB broth with 30 μg/ml kanamycin in New Brunswick Scientific fermentor (New Brunswick, NJ) and the culture was grown at 37°C until the A_{600} was 0.6. Then recombinant protein expression was induced by addition of 1.0 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG), and growth was continued for another 3.0 h. The cells were pelleted, resuspended in 1x binding buffer (20 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 5 mM immidazole) containing 0.1% NP4O (Sigma) and 0.1 mg/mi lysozyme on ice for 30 min, and sonicated briefly. Upon centrifugation at 15000 x g for 15 min, the resulting insoluble pellet was dissolved in 6 M urea in lx binding buffer for 1 h on ice. The urea dissolved supernatant obtained at 15000 x g for 15 min was loaded to an Ni^{2+} activated His-Binding resin column (Novagen, WI) following manufacturer's protocol, and the recombinant protein was eluted with 300 mM immidazole in 1x binding buffer containing 6 M urea. The affinity purified recombinant protein was further purified by preparative SDS-PAGE to remove some lower molecular weight breakdown products of the full length form.

6.1.8 Production of antisera to recombinant FSP95 and immunoblot analysis

Female Lewis rats were immunized with 200 μg of gel purified recombinant protein in Freund' s Complete adjuvant, boosted twice at intervals of two weeks with 200 μg of protein in incomplete Freund's adjuvant, and bled 1 week after each boost. Specificity of the rat antisera for FSP95 was tested by Western blotting against rFSP95 as well as human sperm protein extracts.

For immunoblotting 3.5 μg of the recombinant protein was subjected to 10% SDS-PAGE, and the protein was electrophoretically transferred to nitrocellulose membrane (41). The membrane was cut into strips (each containing - 150 ng), blocked with 5% dry fat-free milk in PBS-Tween (10 mM PBS, pH 7.4, 0.05% Tween 20) for 1 h and incubated with rat antisera at 1:5000 dilution in the blocking buffer for 1 h. Immunodetection was performed with horseradish peroxidase conjugated goat anti-rat IgG (Jackson ImmunoResearch, PA) at 1:5000 dilution in blocking buffer and visualized with a chromogenic substrate, diaminobenzidine (Sigma) in H_20_2 (47). For immunoblotting of sperm proteins, percoll washed sperm were solubilized in lysis buffer containing urea and OBG (see sperm preparation) and subjected to 2-D SDS-PAGE analysis (35). Following electrophoretic transfer of the proteins to nitrocellulose, the membrane was blocked and probed with the primary and secondary antibodies as above.

6.1.9 Micro sperm immobilization assay

Fresh ejaculates were diluted 1:2 with BWW medium and centrifuged at 450 xg for 5 min. The resulting pellet was washed twice in BWW containing human serum albumin (HSA) (10 mg/ml). The pellet was layered with BWW containing HSA (30 mg/mi) for 1 h at 37°C in 5% CO_2. An aliquot of swim-up cells, containing about 98% motile sperm, was diluted in the swim-up medium to 20 million/ml as a working stock. For testing in the immobilization assay, rat antisera against the rFSP95 were diluted 1:1 with swim-up medium and decomplemented at 56°C for 30 min before use. Guinea pig serum which was used as the source of complement was absorbed twice with sperm for 30 min at 4°C. The micro sperm immobilization assay (Isojima and Koyama, 1989 Arch. Androl. 23:185-199) was performed after modification by incubating 10 μl of 1:1 diluted and decomplemented antiserum, 1 μl of the 20 million/ml swim-up sperm and 2 μl of the 1:1 diluted guinea-pig complement (containing ~ 10 to 15 CH5O units/reaction) in a micro PCR tube for 60 min at 37°C. The percentage of motile sperm was determined microscopically in a counting chamber (Humagen Fertility Diagnostics, VA) at 200x. Heat inactivated complement in each test served as a control. The experiment was performed along with CD59 (BioSource International, CA) and SAGA-I monoclonal antibodies (which are known to recognize surface antigens) as positive controls (Jiang and Pillai, 1998, Am. J. Reprod.

Immunol. 39:243-248, Diekman et al., 1997, Biol. Reprod. 57:1136-1144) and with a monoclonal antibody (MHS-10) that recognizes an acrosomal sperm antigen, SP10, as a negative control (Foster et al., 1994, Biol. Reprod. 51:1222-1231). The sperm immobilization value (SIV) was calculated by dividing the percentage of motile sperm in the control (inactivated complement) by that in the test sera with active complement. When the SIV is 2 or more, the test serum is judged as positive for sperm immobilizing antibodies.

6.1.10 Immunofluorescence localization of FSP95 in live or permeabilized sperm

Swim-up spermatozoa were prepared by layering 0.5 ml of semen below 2 ml of BWW medium (Irvine Scientific, CA) containing 3 mg/ml HSA for 1.5 h at 37°C in 5% CO_2. The swim-up cells were then washed and allowed to capacitate in BWW medium containing 30 mg/ml HSA at 37°C in 5% CO_2 for 6 h. The motile capacitated sperm were air dried onto poly-l-lysine coated slides (Polysciences, PA), permeabilized in methanol at -20°C for 30 min, air dried and blocked in 10% normal goat serum in PBS-Tween (10 mM phosphate buffer saline with 0.01% Tween 20) for 30 min. The sperm were then incubated with a 1:100 dilution of the primary antibody (rat anti rFSP95) in the blocking buffer for 2 h at 37°C followed by incubation with the fluorescence (Cyanine, Cy3) conjugated anti-rat (IgG) secondary antibody (Zymed Lab, CA) at 1:100 dilution in the blocking buffer for 1 h at 37°C. The slides were washed in PBS, coated with slow fade (Molecular Probes, OR), and mounted with coverslips. Images were obtained with in an Axioplan fluorescence microscope (Zeiss) at 1000x. For indirect immunofluorescence of live sperm, capacitated sperm were incubated with the primary antibody at a 1:100 dilution in the capacitation medium for 2 h at 37°C in 5% $C0_2$, and subsequently were treated as above.

6.1.11 Electron microscopic immuno localization of FSP95

Pooled sperm were washed twice in Ham's F10 (Gibco BRL, CA) containing 3% sucrose (wash buffer) and were fixed in 4% paraformaldehyde and 0.2% gluteraldehyde in wash buffer for 15 min at 22°C. After removing the fixative by three changes of wash buffer, cells were dehydrated through a series of graded ethanols from 40% to 100%. and embedded in Lowicryl K4M (Electronmicroscopy Sciences, PA). The blocks were polymerized with UV light for 72 h at -20°C and ultrathin sections (100 nm thick) were made.

Immunostaining of thin sections was modified from Berryman and Rodewald, 1990, J. Histochem. Cytochem. 38:159-170). The sections were blocked in undiluted normal goat serum for 15 min at 22°C, and then incubated for 16 h at 4°C with either rat anti rFSP95 or the pre-immune sera from the same animal at 1:50 dilution in wash

buffer containing 1% normal goat serum, 1% BSA and 0.1% Tween 20. After rinsing in wash buffer four times, the sections were incubated with a 1:100 dilution of 5 nm gold-conjugated goat anti-rat IgG (Goldmark Biologicals, NJ) for 1.5 h at 22°C. The specimens were washed in distilled water and stained with uranyl acetate before examination with a JEOL 100CX electron microscope.

6.2 RESULTS

In order to identify human sperm proteins which undergo tyrosine phosphorylation during in vitro capacitation, actively motile sperm obtained after swim-up were incubated for 6 hours under conditions known to allow capacitation to proceed (Yanagimachi, 1994, in: *The Physiology of Reproduction* (Knobil, E., and Neill, J. D., eds) pp. 189-3 17, Raven Press, New York, NY). Extracts containing sperm proteins were separated by 2-D SDS-PAGE and visualized by silver staining (Fig. 1A). Tyrosine phosphorylated proteins were localized on 2-D immunoblots by reaction with the antiphosphotyrosine monoclonal antibody RC20. Fresh, uncapacitated sperm demonstrated a subset of proteins with a low level of phosphorylation on tyrosine residues (Fig. 1B) compared with sperm that were capacitated for 6h (Fig. 1C) and showed increased immunoreactivity of several protein groups ranging in size from 95 to 55 kDa and pl from 4.0 to 5.7. A prominent 95 kDa protein with a pI of ~ 5.3 was seen to be one of the major acidic high molecular weight tyrosine phosphorylated proteins of human sperm (Fig. 1A and 1C). The reaction of antiphosphotyrosine antibody was considered to be specific because preincubation of the RC20 monoclonal antibody with 20 mM phosphotyrosine abolished the immunoreactivity (data not shown). Furthermore tyrosine phosphorylation was almost completely abolished when capacitated cells were incubated with a tyrosine kinase inhibitor, genistein (Fig. 1D).

6.2.1 FSP95 as an auto- and iso-antigen

Immunoblotting of human sperm proteins with sera from infertile male (Fig. 2A) and female (Fig. 2B) subjects previously screened for ASA revealed strong immunoreactivity to a group of sperm proteins with a molecular weight of - 95 kDa. Six serum samples out of 15 from infertile males (40%) and 2 of 6 from infertile females bound the 95 kDa protein group. Among 5 fertile subjects of each sex, 2 male and 1 female sera showed only weak immunoreactivity to the 95 kDa proteins (Fig. 2C and 2D). These findings demonstrated that the 95 kDa proteins were isoantigenic in some women and autoantigenic in some men.

6.2.2 Microsequencing of the 95 kDa antigen by mass spectrometry

To obtain structural information on the identity of one of the 95 kDa tyrosine phosphorylated antigens, microsequencing of a Coomassie stained 2-D SDS-PAGE protein spot was undertaken. The location of the spot cored from the 2-D gel is indicated by a white circle in Fig. 1A. Because of the low amount of the protein available in a well resolved 2-D gel, amino acid sequencing was performed by tandem mass spectrometry on peptides generated by overnight trypsin digestion at 37°C of the protein spot within pieces of the gel. The extracted peptides were concentrated and analyzed by capillary column liquid chromatography electrospray- tandem mass spectrometry. A total of 18 peptide sequences were obtained (Table 1). Database searches using both the molecular weight information (mass mapping by MS-Fit) and sequence information by using Fasta (Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444-2448), allowing a 15% gel derived protein mass tolerance and a 1 Da peptide mass tolerance, did not identify any known protein.

6.2.3 Cloning and sequence analysis of FSP95

To isolate the cDNA encoding the 95 kDa tyrosine phosphorylated protein, a completely degenerate inosine containing forward primer designed from peptide number 6 (Table 1) was used to amplify a 1.0 kb piece of cDNA by 3' RACE PCR from human testicular Marathon ready cDNA (Clontech, CA). The 5' cDNA including an untranslated region was also cloned by PCR producing a 2.1 kb cDNA with a 118 bp overlap at the 3' end. The nucleotide sequence of the full-length cDNA (Fig. 3) consists of 2942 bp with an in-frame start codon at nucleotides 163-165 conforming to a Kozak consensus for the translation initiation site (Kozak, 1991, J. Biol. Chem. 266:19867-19870). The translation start site was further authenticated by the presence of two in frame stop codons at 45 bp and 72 bp upstream of the first ATG sequence. The cDNA contained a 2559 bp open reading frame with untranslated regions of 162 bp at the 5' end, 218 bp at the 3' end, and a polyadenylation signal (ATTAAA) (Juretic and Theus, 1991, FEBS Lett. 290:4-8) 11 bp upstream from the poly(A) tail. The open reading frame encodes a protein of 853 amino acids with a predicted molecular weight of 94.6 kDa and a pI of 6.0. All of the 18 tryptic peptides obtained by microsequencing the 95 kDa protein spot were recovered in the predicted amino acid sequence of the molecule (Fig. 3, underlines), validating that the protein originally identified and cored from the gel had been cloned.

Analysis of the predicted amino acid sequence of FSP95 revealed no N-terminal eukaryotic secretory signal peptide cleavage site (Nielsen et al., 1997, Protein Eng. 10:1-6). Comparison of the protein sequence with the Prosite database (Bairoch et al.,

1997, Nucleic Acids Res. 25:217-221) demonstrated the presence of five potential N-linked glycosylation sites (amino acid 87, 117, 180, 502, 763), fifteen potential casein kinase II phosphorylation sites (amino acid 21, 34, 52, 102, 109, 120, 223, 280, 440, 448, 549, 659, 691, 713, 816), eleven possible protein kinase C phosphorylation sites (amino acid 2, 89, 102, 217, 223, 236, 303, 367, 408, 484, 597), eight myristoylation sites (amino acid 68, 116, 346, 366, 648, 722, 724, 814), and one tyrosine kinase phosphorylation site at amino acid 435. Four potential O-linked glycosylation sites were also found at amino acid numbers 168, 504, 557 and 745 (Hansen et al., 1997, Nucleic Acids Res. 25:278-282). The deduced sequence of FSP95 revealed no apparent transmembrane regions.

Comparison of the deduced FSP95 sequence to the GenBank data base using BLAST (Altschul et al., 1990, J. Mol. Biol. 215:403-410) and FASTA (Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444-2448) revealed that the human sperm FSP95 had closest amino acid identity to a mouse sperm fibrous sheath AKAP, precursor of mouse AKAP82 (pro-mAKAP82) (identity: 33.6%; similarity: 42.5%), and to a human sperm fibrous sheath AKAP, precursor of human AKAP82 (pro-hAKAP82) (identity: 32.4%; similarity: 39.4%) (Turner et al., 1998, J. Biol. Chem. 273:32135-32141, Carrera et al., 1994, Dev. Biol. 165:272-284). The amino acid sequence alignment of these proteins revealed that the middle and part of the N-terminal and C-terminal regions of these molecules contained conserved domains (Fig. 4). The overall identities between FSP95 and these AKAPs suggest that the human sperm FSP95 is encoded by a previously unreported gene. Interestingly, the two potential intracellular anchoring domains of mouse sperm pro-mAKAP82 (Carrera et al., 1994, Dev. Biol. 165:272-284) were conserved in both the pro-hAKAP82 (Turner et al., 1998, J. Biol. Chem. 273:32135-32141) and the human sperm FSP95 cDNAs. Both these intracellular targeting regions were located in the highly basic N-terminal region while the C-termini of these proteins were more acidic, containing residues with long aliphatic side chains. However, the RU-binding domains of human and mouse AKAP82 were not conserved in FSP95.

6.2.4 FSP95 mRNA expression in human tissues

Northern blot (Fig. 5A) and dot blot (Fig. 5B) hybridization were used to analyze the expression of FSP95 mRNA in different human tissues. Northern blots were hybridized with either of two FSP95 probes, a 2337 bp 5' cDNA and a 1.0 kb 3' cDNA. Both probes hybridized to an mRNA of ~ 3.0 kb present only in testis. A broader screening was performed using the 2337 bp 5' cDNA as a probe to hybridize a RNA dot blot containing poly(A)+ RNAs from 50 different human tissues. Among the tissues examined,

a strong hybridization signal was observed only from testis mRNA, indicating a testis specific expression pattern of human sperm FSP95 gene.

6.2.5 Expression of rFSP95 and Western blot analysis

The cDNA sequence encoding the FSP95 from residues 1 to 779 was cloned into the bacterial expression vector pET28b. When the bacteria were induced with LPTG for 1.5 or 3 h, rFSP95 with molecular weight of - 97 kDa was produced (Fig. 6A). The recombinant protein was purified by utilizing the high affinity of the six-His domain for Ni^{2+} ions immobilized on Sepharose (Van Dyke et al., 1992, Gene 111:99-104). The affinity purified recombinant FSP95 was further purified by preparative gel electrophoresis (Fig. 6B) and used as an immunogen to inoculate rats. The rat antibody to rFSP95 recognized both the rFSP95 and the "native" FSP95 present in sperm extracts and stained the 95 kDa spot at a pI ~ 5.3 which was originally microsequenced (Fig. 6C and 6D). Pre-immune rat sera did not react with either rFSP95 or with sperm proteins.

6.2.6 Localization of FSP95 in human spermatozoa

The possibility that FSP95 is exposed on the plasmalemma of fresh non capacitated human spermatozoa was evaluated first by a modified micro sperm immobilization assay (Isojima and Koyama, 1989, Arch. Androl. 23:185-199). Immobilization of fresh sperm was not observed in presence of rat antiserum against the rFSP95 and guinea pig complement (SIV = 0.91 ± 0.12 mean ± SD; n=6). Furthermore, live capacitated sperm did not show immunofluorescent staining over any domain (data not shown). Together these results indicate a lack of FSP95 at the cell surface.

Indirect immunofluorescence analysis of capacitated and methanol permeabilized ejaculated human spermatozoa using rat serum against rFSP95 localized FSP95 to the entire length of the principal piece of the flagellum (Fig. 7B) of 100 % sperm. The head, midpiece and the end piece remained unreactive (Fig. 7A and 7B). Pre-immune antisera showed no immunofluorescence (Fig. 7C and 7D).

The intracellular distribution of FSP95 was examined at the ultrastructural level by post-embedding immunolabelling of ultrathin sections of washed ejaculated human spermatozoa. Gold particles indicating the localization of FSP95 were associated with the entire thickness of the fibrous sheath in both longitudinal and cross sections. The gold particles were observed over the circumferential ribs while the central zone of the longitudinal columns (LC) remained unstained (Fig. 8A and 8B, indicated by arrows). No label was associated with the outer dense fibers or the axoneme. Only a rare gold particle was observed in sections exposed to preimmune control rat sera (Fig. 8C and 8D).

6.1.7 FSP95 is a capacitation induced tyrosine phosphorylated protein

To provide further evidence that the protein which had been selected, microsequenced and cloned did indeed undergo tyrosine phosphorylation during in vitro capacitation, the following experiment was performed. Protein extracts from both capacitated and non-capacitated sperm were resolved by high resolution (23x23 cm) 2-D SDS-PAGE, electroblotted, and analyzed with rat antisera against rFSP95 and antiphosphotyrosine antibody RC20. Immunoreactive forms of FSP95 with a pI of ~ 5.3, evident in extracts of non-capacitated sperm (Fig. 9A), became less abundant in capacitated sperm (Fig. 9B). Concomitantly, new immunoreactive forms of FSP95 with more acidic pIs appeared. This shift in charge towards acidic FSP95 isoforms was accompanied by increased tyrosine phosphorylation after capacitation (compare Fig. 9C and 9D) and the appearance of more acidic FSP95 isoforms which were phosphorylated. This demonstration using antibody to rFSP95 provides formal proof that the FSP95 which was cloned and sequenced undergoes tyrosine phosphorylation during capacitation and has isoforms which become more acidic in the process.

6.3 DISCUSSION

In this report we describe the cloning and characterization of a novel human sperm antigen, designated FSP95 in consideration of its fibrous sheath localization and calculated molecular weight of 95 kDa. In order to characterize FSP95 from a 2-D SDS-PAGE protein spot derived microsequence, the 3' cDNA of the molecule was amplified using a single gene-specific inosine containing primer and an adapter primer from human testicular Marathon cDNA. The obtained cDNA revealed the presence of 4 FSP95 microsequenced peptides embedded in its open reading frame. The 5' end of the cDNA was then similarly cloned by 5' RACE revealing two in frame stop codons upstream of the most 5' methionine identified. This indicates that the reported cDNA represents the complete coding sequence of FSP95. The translation start site also conforms to the essential Kozak consensus sequence. Embedded within the deduced amino acid sequence are all 18 microsequenced peptides obtained from the FSP95 protein spot (Table 1; Fig. 3) validating that the protein spot originally cored was cloned. The success of this cloning strategy using a single microsequence derived primer is important because it was completed in only about three weeks. Therefore, this approach may be an important avenue in converting microsequence data into complete sequence information and highlights the increasing importance of proteome based cloning.

From a methodological view it is important to note that the peptide sequence used to initiate the cloning experiment (peptide 6, Table 1) was confirmed by combining

interpretation of the CAD spectrum of the peptide in the digest, interpretation of the CAD spectrum of the N-terminal derivatized peptide (Cardenas et al., 1997, Rapid Commun. Mass Spectrom. 11:1271-1278) and comparison of CAD spectra of the digest peptide with a synthetic peptide. Minor discrepancies were observed in comparing the amino acid sequences derived from the mass spectrometry data and the cDNA data for peptides 6, 12 and 18 due to changes in a single base. The differences produce an F to S substitution in peptide 6, either an E to V or a T to A substitution in peptide 12 and an E to G substitution in peptide 18. These substitutions could be due to an inter-individual DNA polymorphism between the donors from which FSP95 was microsequenced and the donor from whom the Clontech Marathon testicular cDNA was constructed.

Characterization of proteins whose patterns of tyrosine phosphorylation changes with capacitation in human spermatozoa is an important step in dissecting the involvement of various signal transduction pathways in capacitation related changes which are preparatory for mammalian fertilization. Using 1-D gel electrophoresis human spermatozoa were previously shown to undergo increased tyrosine phosphorylation of two major high molecular weight proteins of ~ 105 and ~ 80 kDa during the process of capacitation (Carrera et al., 1996, Dev. Biol. 180:284-296, Leclerc et al., 1998, J. Androl. 19:434-443, Luconi et al., 1995, Mol. Cell Endocrinol. 108:35-42). In the present study, 2-D Western analysis comparing capacitated, uncapacitated and genistein treated sperm revealed multiple high molecular weight proteins which showed increased tyrosine phosphorylation during capacitation, including the antigen FSP95 (Fig. 2C). Based upon its abundance on phosphotyrosine containing blots, FSP95 emerges as one of the major acidic high molecular weight tyrosine phosphorylated proteins of capacitated human spermatozoa thus far described. Indeed, a potential tyrosine kinase phosphorylation site was identified in FSP95 at residue number 435. The change in pI of FSP95 toward acidic isoforms during in vitro capacitation as examined with antibodies to phosphotyrosine residues and to recombinant FSP95 (Fig. 9) provide formal proof that FSP95 is a tyrosine kinase substrate.

The inhibition of tyrosine phosphorylation in capacitated sperm by genistein (Fig. 2D) suggests that a tyrosine kinase is active and present in human sperm. The presence of a tyrosine kinase has previously been demonstrated in ejaculated human spermatozoa and in the mid-piece tail region of boar spermatozoa using antibodies against tyrosine kinase purified from boar male germ cells (Berruti and Borgonovo, 1996, J. Cell. Sci. 109:851-858). The decrease in protein phosphorylation of tyrosine residue observed in capacitated sperm in presence of genistein below that of the uncapacitated cells (Fig. 1) is likely due to the presence of phosphatases, perhaps activity of the calmodulin-dependent protein phosphatase calcineurin (Tash et al., 1988, J. Cell Biol. 106:1625-1633). Such an

interpretation is supported by the observation that Ca^{++} induced dephosphorylation is calmodulin-calcineurin dependent (Carrera et al., 1996, Dev. Biol. 180:284-296). As tyrosine phosphorylation of mammalian spermatozoa is associated with sperm hyperactivation (Visconti and Kopf, 1998, Biol. Reprod. 59:1-6), the capacitation induced tyrosine phosphorylation of FSP95 could play a role in sperm motility. Furthermore, availability of a major tyrosine kinase substrate may provide a means to characterize the participating tyrosine kinase(s) involved in these sperm functions.

Comparison of the deduced amino acid sequence of the FSP95 cDNA with the available database reveals its highest homology with mouse and human sperm pro-AKAP82 (34% and 32% amino acid identity, respectively), which are protein kinase A anchor proteins that sequester PK-A to subcellular locations. This suggests that FSP95 with these AKAPs reveals that FSP95 possesses two potential intracellular targeting domains (Fig. 4). Both these domains lie within an N-terminal basic region similar to the AKAPs 75 and 79 (Glantz et al., 1993, J. Biol. Chem. 268:12796-12804, Carr et al., 1992, J. Biol. Chem. 267:6816-16823). Interestingly, however, the predicted RII-binding domain of these sperm AKAPs was found to be lacking in FSP95 as judged by poor conservation and lack of the putative amphipathic helix binding motif in this region (Carr et al., 1991, J. Biol. Chem. 266:14188-14192). In a recent study, the presence of an AKAP of approximately 110 kDa and the regulatory subunits of PK-A (RIIα, RIIβ and RIβ) has been demonstrated in human, bovine and monkey spermatozoa (Vijayaraghavan et al., 1997, J. Biol. Chem. 272:4747-4752). Moreover, it has been suggested that the anchoring of the regulatory subunity of PK-A to bovine sperm AKAP, independent of PK-A catalytic activity, is essential for the regulation of sperm motility (Vijayaraghavan et al., 1997, J. Biol. Chem. 272:4747-4752). Therefore, considering the unique testis specific expression pattern of FSP95 (Fig. 5), and its similarity to sperm AKAPs FSP95 could be explored for the possible development of a cell permeable anchoring inhibitor peptide for the formulation of a topical spermiostatic agent for human. Furthermore, the identification of FSP95 as a tyrosine kinase substrate and its similarity to sperm AKAP may suggest possible interrelationship between PK-A and tyrosine kinase signaling pathways, because tyrosine phosphorylation and capacitation in other mammals has been shown to be upregulated by a cAMP/PK-A dependent pathway (Visconti and Kopf, 1998, Biol. Reprod. 59:1-6, Galantino-Homer et al., 1997, Biol. Reprod. 56:707-719, Visconti et al., 1995, Development 121:1139-1150).

Since it is recognized by sera with antisperm antibodies from both men and women, FSP95 is both- iso- and auto-antigenic. This finding is in concert with the remarkable testis-specificity observed for expression of the FSP95 transcript. Many sperm

proteins are antigenic in nature because they are tissue specific and do not appear until puberty when meiosis is initiated and sperm specific genes begin to be transcribed (Tung et al., 1985, in: *The Autoimmune Diseases* (Rose and MacKay, eds.) pp. 537-590, Academic Press, New York, NY). During the induction of self tolerance during the neonatal period such neo-antigens are probably not recognized by the immune system (Tung et al., 1985, in: *The Autoimmune Diseases* (Rose and MacKay, eds.) pp. 537-590, Academic Press, New York, NY) while after meiosis the later stages of male germ cells are sequestered from the immune system by the blood-testis barrier and the blood-epididymis barrier (Setchell et al., 1990, J. Reprod. Immunol. 18:19-32, Hoffe and Hinton, 1984 Biol. Reprod. 30:991-1004). Although several human sperm-specific auto- and iso-antigens are recognized by infertile patient sera (*e.g.*, SAGA 1, SPAG2, 5P17, FAI,) (Diekman et al., 1997, Biol. Reprod. 57:1136-1144 50, Diekman et al., 1998, Mol. Reprod. Dev. 50:284-293, Lea et al., 1997, Fertil. Steril. 67:355-361, Naz et al., 1984, Science 225:342-344), only a very few have been found on the spermatozoal surface (e.g., SAGA1) (Diekman et al., 1997, Biol. Reprod. 57:1136-1144) - a critical criteria for the selection of a potential contraceptive immunogen (Herr, 1996, in: *Contraceptive Research and Development* (Harrison and Rosenfield, eds.) pp. 401-429, National Academy Press, Washington DC). Thus, FSP95 is not a candidate contraceptive vaccinogen because both the micro sperm immobilization assay and immunofluorescence indicated that it is not on the sperm surface. Instead, immunocytochemistry using antisera against rFSP95 localized FSP95 to the cytoplasm of the principal piece of the tail.

Immunoelectronmicroscopy identified the antigen in association with the ribs of the fibrous sheath (Fig. 8), which are believed to be involved in defining the shape of the flagellar beat (Eddy and O'Brien, 1994, in: *The Physiology of Reproduction,* (Knobil and Neill, eds.) Vol. 1, pp. 29-77, Raven Press, New York, NY). In demembranated mouse spermatozoa, it has been demonstrated that sliding of the fibrous sheath towards the head is accompanied by extrusion of microtubules towards the distal end (Si and Oknno, 1993, Exp. Cell Res. 208:170-174). The sliding of the fibrous sheath was cAMP dependent and the sliding velocity depended on the ATP concentration. Although several proteins have been identified on gels of isolated human sperm fibrous sheath (MW: 97, 76, 62, 55, 33, 28, 25 kDa) (Jassim et al., 1992, Hum. Reprod. 7:86-94), to our knowledge only the human sperm fibrous sheath protein, hAKAP82 has been characterized at the molecular level (Turner et al., 1998, J. Biol. Chem. 273:32135-32141), other than FSP95. The recently reported human testis specific gene termed 'hi' encoding a predicted protein although showed 92% identity to the human sperm fibrous sheath protein pro-hAKAP82 and 27%

identity with FSP95, however, its localization in spermatozoa has not been demonstrated (Mohapatra et al., 1998, Biochem. Biophys. Res. Commun. 244:540-545).

Anomalies in fibrous sheath structure, particularly disorganization of components and asymmetrical location of the longitudinal columns have been shown to be associated with severe sperm immobility related to sterility and flagella dyskinesia respectively (Chemes et al., 1987, Fertil. Steril. 48:664-669, Serres et al., 1986, Cell Motil. Cytoskeleton 6:68-76). Although the molecular basis of these defects is presently unknown, mutations of the genes encoding fibrous sheath proteins may be related to the disorganization of fibrous sheath structure leading to flagellar dyskinesia. Further studies on disruption of the FSP95 gene would clarify its role in fibrous sheath organization as well as aid in evaluating the importance of tyrosine phosphorylation of FSP95 in hyperactivation and capacitation.

The invention is not to be limited in scope by the specific embodiments described which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

All references cited herein are incorporated herein by reference for all purposes.